

Molecular and Genetic Studies of the *uncoordinated* Gene of  
*Drosophila melanogaster*.

A thesis submitted for the degree of  
Doctor of Philosophy  
of The Australian National University

by

Marion Joy Healy

June 1985



# DECLARATION

The research described in this thesis is entirely my own work, except as indicated below:

Dr. Robyn Russell constructed the  $\lambda$  EMBL 3a Canton-S library, and together with Dr. George Miklos, performed the bulk of the library screening experiments (Section 3.1.1, Table 3.2); Dr. Ian Boussy prepared the polytene chromosomes (Section 5.1.2); and Julie Higginbotham assisted in the construction of the homozygous deficiency genotypes (Section 4.5.1) and the initial collection of the wildtype developmental stages (Section 4.1).

Marion J. Healy.

Marion J. Healy



ABSTRACT

The technique of 'chromosome walking' was used to isolate approximately 60kbp of DNA from the *uncoordinated* complementation group, located in a region spanning the junction between the euchromatin and heterochromatin of the X chromosome of *Drosophila melanogaster*. Characterization of the cloned DNA revealed that it could be divided into two distinct regions. The first contains sequences that are unique and are well conserved between strains. The remainder are characterized by units repeated in tandem arrays and separated by a member of a high copy number sequence family, and which are polymorphic between, and within, strains.

Transcriptional analysis revealed that two, and possibly as many as six, transcripts are generated from the *uncoordinated* complementation group, one of which is non-adenylated. The polyadenylated transcripts are generated in a temporally specific manner whereas the nonadenylated species is present throughout the organism's life cycle. There is also spatial differentiation in the production of the polyadenylated and non-adenylated RNA species found in the adult. It is proposed that the multiple transcripts are generated by differential processing of a primary transcript.

Confirmation that the isolated DNA derived from the *uncoordinated* gene was obtained from characterizing the molecular basis of the mutation in seven *uncoordinated* alleles. Analysis of transcripts from these alleles confirmed that the adult polyadenylated transcript encoded at least part of the *uncoordinated* gene product, although the role of the nonadenylated transcript is still unclear.

Comparison of the *uncoordinated* complementation group with other cloned genes from *Drosophila* suggests that its molecular organization is characteristic of genes located near, or at, regions of intercalary heterochromatin and is consistent with the proposal that intercalary

heterochromatin is at least partly composed of repeated sequences. The arrangement of DNA sequences in subdivisions 19E/F was shown to be heterogeneous by comparing the region cloned in this study, and other cloned sequences in division 19F, with well characterized euchromatic regions. The sequence arrangement in 19E/F is consistent with the model that the euchromatin-heterochromatin transition zone consists of unique coding regions interspersed with repetitive elements of variable length.

Properties of the *uncoordinated* transcription unit have many similarities to those of genes located at other sites in the genome, indicating that such properties are not involved in the differential regulation of genes in distinctive chromosomal regions. However, some of the properties of the *uncoordinated* transcription unit are similar to those of genes whose products have been identified in neural tissue. This raises the interesting possibility that the *uncoordinated* gene product has a neurological function.

ACKNOWLEDGEMENTS

I am extremely grateful to the two principal supervisors of my Ph.D. programme: Dr George Miklos, for stimulating my enthusiasm for molecular developmental biology; and Dr. Robyn Russell, for sharing her scientific rigour and considerable technical skills.

For their generous assistance, I am especially grateful to Valmai Hicks, Julie Higginbotham and Paul Simpson. I thank Dr. John Oakeshott and Rick Tearle for enlightening discussions and Dr. John Watson for encouragement. I also thank the members, and former members of the laboratory, Veronica Ross, Bronwyn Gray and Dr. Jane Davies, for their friendship and support.

I sincerely thank those people who assisted me in the preparation of this thesis: Garry Brown for the excellent drawings, the R.S.B.S. Photography Unit, Drs. Tony Howells and Rob Saint for reading the thesis, Barbara Piper for her excellent typing, Dave Willcocks for exerting his considerable influence and Phil Anderson for guiding me through use of the word processor. Most of all, I thank my husband, Michael, for encouragement and tolerance.

The financial support of a Commonwealth Postgraduate Research Award is gratefully acknowledged.

ABBREVIATIONS

|                     |                                      |
|---------------------|--------------------------------------|
| <i>Adh</i>          | alcohol dehydrogenase gene           |
| Ap                  | ampicillin                           |
| bp                  | base pairs                           |
| Cm                  | chloroamphenicol                     |
| DEPC                | diethylpyrocarbonate                 |
| DMSO                | dimethylsulphoxide                   |
| DTT                 | dithiothreitol                       |
| kbp                 | kilobase pairs                       |
| p.f.u.              | plaque forming unit                  |
| poly A <sup>+</sup> | polyadenylated                       |
| Tc                  | tetracycline                         |
| TEMED               | N,N,N',N'-tetramethylethylenediamine |
| <i>unc</i>          | uncoordinated gene                   |
| U.V.                | ultra-violet                         |

Other abbreviations are in common usage and have been previously described (*J. Biol. Chem.* 241: 527-533, 1966).

## TABLE OF CONTENTS

|   | Page |
|---|------|
| <u>CHAPTER ONE</u>  |      |
| <u>INTRODUCTION</u>   | 1    |
| 1.1 Drosophila Heterochromatin  | 2    |
| 1.1.1 Cytological and genetic analysis of heterochromatin   | 2    |
| 1.1.2 Molecular analysis of heterochromatin   | 5    |
| 1.1.3 Intercalary heterochromatin   | 8    |
| 1.2 Chromosome Position and Gene Function   | 9    |
| 1.2.1 Position effect and variegation   | 9    |
| 1.2.2 Transformation studies  | 12   |
| 1.3 Molecular Analysis of the Structural Organization and Regulation of Genes from <i>D. melanogaster</i> | 14   |
| 1.3.1 Structural organization of cloned genes   | 14   |
| 1.3.2 Regulation of gene expression   | 17   |
| 1.4 The Proximal Region of the X-Chromosome: A Molecular and Genetic Analysis                             | 23   |
| <u>CHAPTER TWO</u>  |      |
| <u>MATERIALS AND METHODS</u>  |      |
| 2.1 Materials   | 26   |
| 2.1.1 Chemicals and reagents  | 26   |
| 2.1.2 Enzymes   | 27   |
| 2.1.3 Bacterial strains and media   | 28   |
| 2.1.4 DNA vectors   | 29   |
| 2.1.5 Ribonuclease free glassware and solutions   | 29   |
| 2.1.6 Drosophila stocks and media   | 29   |
| 2.2 Nucleic Acid Preparative Procedures   | 30   |
| 2.2.1 Genomic DNA   | 30   |
| 2.2.2 Supercoiled plasmid DNA   | 30   |
| 2.2.3 Bacteriophage DNA   | 31   |
| 2.2.4 Total cellular RNA  | 33   |
| 2.2.5 Polyadenylated RNA  | 34   |
| 2.3 Preparation of Nucleic Acids for Electrophoresis  | 35   |
| 2.3.1 Preparation of DNA samples  | 35   |
| 2.3.2 Preparation of RNA samples  | 35   |
| 2.4 Electrophoresis of Nucleic Acids  | 36   |
| 2.4.1 Agarose gel electrophoresis of DNA  | 36   |
| 2.4.2 Polyacrylamide gel electrophoresis of DNA   | 37   |
| 2.4.3 Agarose gel electrophoresis of RNA  | 37   |

|   | Page |
|---|------|
| 2.5 Molecular Cloning Protocol                  | 38   |
| 2.5.1 Preparation of DNA samples                | 38   |
| 2.5.2 Dephosphorylation of vectors              | 38   |
| 2.5.3 Ligation condition                        | 38   |
| 2.5.4 Transformation procedures                 | 39   |
| 2.5.5 Selection of recombinants                 | 39   |
| 2.6 Screening Genomic Libraries                 | 40   |
| 2.7 Hybridization Procedures                    | 41   |
| 2.7.1 Preparation of hybridization probes       | 41   |
| 2.7.2 Southern blot hybridizations              | 43   |
| 2.7.3 Northern blot hybridizations              | 45   |
| 2.8 Drosophila Techniques                       | 46   |
| 2.8.1 Drosophila culture                        | 46   |
| 2.8.2 Collection of life stages                 | 46   |
| 2.8.3 Preparation of salivary gland chromosomes | 47   |

### CHAPTER THREE

|  |    |
|--|----|
| <u>ISOLATION AND ANALYSIS OF GENOMIC DNA FROM THE UNCOORDINATED<br/>COMPLEMENTATION GROUP</u>  | 49 |
| 3.1 Isolation of Genomic DNA from the <i>uncoordinated</i><br>Complementation Group  | 50 |
| 3.1.1 Isolation of lambda clones by chromosome walking   | 50 |
| 3.1.2 Determination of the restriction enzyme cleavage<br>maps of lambda clones  | 51 |
| 3.2 Molecular Analysis of Genomic DNA from the <i>uncoordinated</i><br>Complementation Group   | 53 |
| 3.2.1 Genomic representation of DNA sequences from the<br><i>uncoordinated</i> complementation group   | 53 |
| 3.2.2 Characterization of high copy number sequences   | 55 |
| 3.3 Interstrain Variability  | 57 |
| 3.3.1 Comparison of the sequences from the <i>uncoordinated</i><br>complementation group in two wildtype strains   | 57 |
| 3.3.2 Comparison of hybridization patterns of repeated<br>sequences from the <i>uncoordinated</i> complementation<br>group in different wildtype strains | 59 |
| 3.4 Correlation of Genetic and Molecular Maps  | 60 |
| 3.5 Conclusions  | 61 |



CHAPTER FOURTRANSCRIPTIONAL ANALYSIS OF THE UNCOORDINATED COMPLEMENTATION GROUP

|       |   |    |
|-------|---|----|
| 4.1   | Identification of Transcripts Homologous to <i>uncoordinated</i> Sequences                        | 63 |
| 4.2   | Temporal Specificity of Transcripts Homologous to <i>uncoordinated</i> Sequences                  | 66 |
| 4.3   | Spatial Distribution of Transcripts Homologous to <i>uncoordinated</i> Sequences                  | 67 |
| 4.4   | Direction of Transcription of RNA Products Homologous to <i>uncoordinated</i> Sequences           | 67 |
| 4.5   | Localization of Transcripts to the <i>uncoordinated</i> Complementation Group                     | 69 |
| 4.5.1 | Construction of genotypes homozygous deficient for the <i>uncoordinated</i> complementation group | 70 |
| 4.5.2 | Localization of adult transcripts   | 70 |
| 4.6   | Conclusions   | 72 |

CHAPTER FIVEANALYSIS OF MUTATIONS IN THE UNCOORDINATED GENE

|       |   |    |
|-------|---|----|
| 5.1   | Analysis of DNA Sequences from <i>uncoordinated</i> Mutant Alleles          | 74 |
| 5.1.1 | Southern blot hybridization analysis of <i>uncoordinated</i> mutant alleles | 74 |
| 5.1.2 | Cytology of <i>uncoordinated</i> mutants                                    | 78 |
| 5.2   | Transcriptional Analysis of <i>uncoordinated</i> Mutant Alleles             | 79 |
| 5.3   | Conclusions   | 80 |

CHAPTER SIXDISCUSSION

|       |  |    |
|-------|--|----|
| 6.1   | Molecular Organization of the <i>uncoordinated</i> Complementation Group | 82 |
| 6.1.1 | The <i>uncoordinated</i> complementation group                           | 82 |
| 6.1.2 | The euchromatin-heterochromatin transition zone                          | 86 |
| 6.2   | Transcription of the <i>uncoordinated</i> Gene                           | 87 |
| 6.2.1 | Structural characteristics of the transcription unit                     | 87 |
| 6.2.2 | Regulation of the <i>uncoordinated</i> transcription unit                | 91 |
| 6.3   | Function of the <i>uncoordinated</i> Gene Product                        | 95 |
| 6.4   | Future Perspectives  | 98 |

|  | Page |
|--|------|
| 6.4.1 Regulation of the <i>uncoordinated</i> complementation group | 98   |
| 6.4.2 Function of the <i>uncoordinated</i> gene                    | 102  |

|                     |     |
|---------------------|-----|
| <u>BIBLIOGRAPHY</u> | 104 |
|---------------------|-----|



## CHAPTER ONE

INTRODUCTION

The studies described in this thesis were undertaken with the aim of using recombinant DNA techniques to analyse a gene in a region of the genome where the chromosomal organization was changing from cytologically determined euchromatin to heterochromatin (Cooper, 1959). Studies of *Drosophila* indicate that when genes are relocated to chromosomal heterochromatin their expression may be altered, suggesting that the regulation of expression is subjected to different and/or additional controls. However, in a wildtype genome, some loci are located in, or adjacent to, such heterochromatic regions, and the question arises as to the organization and regulation of unique coding sequences interspersed with the heterochromatic segments. It is possible that all genes are subject to the same regulatory mechanisms, and the factors controlling compaction of heterochromatin do not affect the adjacent genes. Alternatively, there may be additional constraints in the organization and regulation of genes adjacent to heterochromatin, which counteract the influence of factors controlling heterochromatinization. It was anticipated that this problem could be approached by examining the structural organization, regulation and flanking sequences of a gene in such a chromosomal region, and by comparing these properties with those of genes located in euchromatic regions. The gene chosen for study was *uncoordinated* (*unc*), located in the proximal region of the X chromosome and formally placed in polytene chromosome band 19E8.

It is now apparent that, in eukaryotic organisms, there are multiple levels of regulation of gene expression (Darnell, 1982). Major levels of control are exerted during the initiation of transcription and the processing

of primary transcripts within the cell nucleus, although cytoplasmic control mechanisms are, at present, relatively unexplored. Differential translation of the mature transcript has been demonstrated (Fruscoloni *et al.*, 1983), and more recently the inhibition of translation by an anti-mRNA has been proposed (Pestka *et al.*, 1984). Post-translational control mechanisms have been described whereby large polypeptides are proteolytically cleaved to yield overlapping peptides (Herbert and Uhler, 1982). In order to initiate an investigation into the mechanisms used to specify its temporal and spatial expression, the *uncoordinated* gene was analysed at the transcriptional level.

A second, although longer term aim of this work, was to determine the function of the *uncoordinated* gene and assess its role in governing the organism's development. One of the central questions in biology asks how a zygote develops in an orderly manner into the highly diverse cells and tissues of a mature organism. This question is of interest to a variety of disciplines such as: the establishment, organization and utilization of positional information in the developing embryo; determination of cell fate; the temporal and tissue specific production of gene products and the diversity of neurons and the precision with which they are interconnected. It was therefore of interest to consider the impact on developmental processes of the *uncoordinated* gene.

## 1.1 Drosophila Heterochromatin

### 1.1.1 Cytological and genetic analysis of heterochromatin

The term 'heterochromatin' (Rieger *et al.* 1968) was introduced to describe regions of the mitotic chromosome which retain a dense and compact structure during interphase, and which differ from euchromatin in the degree to which they change from telophase to interphase. It has been demonstrated that the condensation of euchromatic and heterochromatic portions of the genome are at

least partly under separate genetic control (Gatti *et al.*, 1983). In *D. melanogaster*, analysis of mitotic chromosomes has revealed that the entire Y chromosome, the proximal half of the X chromosome (Figure 1.1(a)) and proximal quarters of each arm of chromosomes 2 and 3 are heterochromatic. Cytologically, the heterochromatin of the X and Y chromosomes is divisible into different regions by the presence of the nucleolus organizer and secondary constrictions (reviewed by Cooper, 1959), and the heterochromatic block on chromosome 2 is also subdivided by a constriction (Kaufman, 1934). Fluorochromes resolve the heterochromatin further into characteristic banding patterns (Holmquist, 1975).

Heterochromatin is defined as either constitutive, whereby both the homologous chromosomal regions are condensed, or facultative, where condensation of two homologous chromosomes differ. Constitutive heterochromatin is regarded as a permanent feature; however Cooper (1959) has proposed that, in response to defined environmental cues and/or developmental signals, heterochromatic regions can decondense and become transcriptionally active. For example, heterochromatin does not appear in the very early embryo and although the Y chromosome is usually regarded as heterochromatic, at some stages of spermatogenesis it clearly is not (reviewed in Brown 1966; Spofford 1976).

There are several hypotheses as to the nature of the junction between euchromatin and heterochromatin. It has been proposed that a definitive junction exists between the cytologically distinguishable regions, and in the case of the X chromosome, this junction is thought to be located proximal to subdivision 20D-E (Schalet and Lefevre, 1976). In contrast, Lifschytz (1978) has proposed that such a distinct border does not exist and that the base of the X chromosome is composed of heterochromatic segments of different lengths that alternate with euchromatic segments. The length of the heterochromatic blocks in the mitotic chromosome could then be inferred from genetic studies.

A notable difference between the heterochromatin and euchromatin of mitotic chromosomes of *Drosophila* is their relative replication during polytenization. During this process, euchromatin replicates to form characteristic chromomeres (or bands), whereas heterochromatin (termed  $\alpha$ -heterochromatin) replicates very little, if at all, and is thought to fuse to form the chromocenter (Figure 1.1(b)). The poorly banded segment between the chromosome arms and the chromocenter is termed  $\beta$ -heterochromatin and this region, together with the chromocenter, undergo replication during polytenization, (Gall, 1973). Consequently, it has proved difficult to delineate, on a polytene chromosome, the start of the heterochromatin as defined on a mitotic chromosome, and its location has been deduced from genetic studies.

Cytogenetic analyses of *Drosophila* chromosomes has revealed that heterochromatin is more sensitive than euchromatin to breakage after radiation treatment (Hannah, 1951). Mukhina *et al.* (1981) summarized the data, and although some of the experiments included in this summary may be misleading because of the difficulty in determining the location of the heterochromatin on the salivary gland polytene chromosome, there is nonetheless convincing evidence that heterochromatin is more susceptible to chromosome breakage (for example, Schalet and Lefevre, 1976). This property has been used to infer the location of the heterochromatin in polytene chromosomes.

There has been considerable debate in the literature as to whether heterochromatic regions of chromosomes contain genetic information. Extensive analysis of the proximal heterochromatin of chromosome 2 (Hilliker and Holm, 1975; Hilliker, 1976) has localized a number of functions to this region and from these studies genic sequences have been estimated to occur at a density of approximately 1% of the occurrence of genetic loci in euchromatin per kilobase pairs (kbp) of DNA (Hilliker *et al.*, 1980). It is still unclear as

to whether any sex-linked genes (i.e. those genes not covered by the Y chromosome) are located in the X chromosome heterochromatin, again because of the uncertainty in the location of the euchromatin-heterochromatin junction in mitotic chromosomes and of heterochromatin in salivary gland chromosomes (Schalet and Lefevre, 1973; 1976). However, a number of functions that are also influenced by the Y chromosome have been mapped to the X heterochromatin, including the *bobbed* locus and regulatory elements affecting the level of endoreplication and expression of genes on other chromosomes (reviewed in Hilliker *et al.*, 1980). The total number of well-defined genes on the Y chromosome is eight, consisting of *bobbed* and the genes coding for fertility factors (reviewed in Williamson, 1976), in addition to several other ill-defined loci common to both the X and Y chromosomes. The heterochromatin of the third and fourth chromosome has not been well characterized with respect to genic sequences.

#### 1.1.2 Molecular analysis of heterochromatin

Highly repeated DNA sequences (satellite DNA) have been localized predominantly to heterochromatin in a variety of species (reviewed in John and Miklos, 1979). Six distinct satellites have been identified in *D. melanogaster*, which together comprise approximately 25% of the genome (Peacock *et al.*, 1977; Brutlag *et al.*, 1977; Brutlag, 1980). Four major satellite species of densities 1.672, 1.686, 1.688 and 1.705g/cc (satellites I, II, III and IV, respectively) have been extensively characterized, whereas the two others, of densities 1.690 and 1.697 g/cc, have been less well studied. Satellites I, II and IV are regarded as simple and consist of a repeating unit of 5 to 10 base pairs organized into tandem arrays. Each of these satellites is not homogeneous, but contains alternative forms. The 1.688 species is more complex having a repeating unit of 359bp (Carlson and Brutlag, 1977) and



restriction enzyme cleavage site variations within the repeating unit (Hsieh and Brutlag, 1979(b)). In addition, Carlson and Brutlag (1979) have identified a variant of this unit, which consists of tandem copies of a novel 254bp unit and is identical to the 359bp repeat except for a 98bp deletion.

*In situ* hybridization techniques have localized the satellite sequences to only a few sites in the genome (Peacock *et al.*, 1977). These sites are predominantly heterochromatic, although satellite sequences have also been found at divisions 21D1,2 and 81. The 1.688 satellite is found exclusively on the X chromosome (Hilliker and Appels, 1982), and although the remaining satellites are not restricted to any one chromosome, the quantitative and qualitative organization of satellite sequences on each chromosome is unique. In some instances, the relative positions of the satellite sequences have been identified (for example, Steffensen *et al.*, 1981).

Comparisons of the distribution of satellite DNA and heterochromatin through-out the genome (Peacock *et al.*, 1977; John and Miklos, 1979) has revealed that satellite DNA cannot be the only component of heterochromatin. Furthermore, the application of molecular cloning techniques to the study of repeated DNA has revealed that sequences of the middle repetitive category, including members of *copia* (Strobel *et al.*, 1979), *copia*-like (Finnegan *et al.*, 1977; Modolell *et al.*, 1983), fold-back (Truett *et al.*, 1981) and I-factor (Bucheton *et al.*, 1984) families and type-1 insertion sequences of rRNA genes (Peacock *et al.*, 1981) and their adjacent elements (Kidd and Glover, 1980; Dawid *et al.*, 1981) are also located in chromocentric heterochromatin. In addition, moderately repeated sequences found associated with the heat-shock genes at subdivision 87C ( $\alpha\beta$ units) (Lis *et al.*, 1981), telomeric-related sequences (Young *et al.*, 1983), short, simple repeated sequences (Simpson, 1984) and clustered, scrambled sequences (Wensink *et al.*, 1974; 1979) have also been found in  $\beta$ -heterochromatin. As sequences homologous to

*copia* and the type 1 insertion element have been found adjacent to cloned segments of satellite DNA (Carlson and Brutlag, 1978(a); 1978(b)), these two middle repetitive sequences are probably located in mitotic heterochromatin.

The data demonstrate convincingly that both satellite and middle repetitive DNA are major components of the heterochromatin of *D. melanogaster*. Two moderately repeated sequence families are interspersed with highly repeated sequences, and the genetic analysis of heterochromatin demonstrates that genetic functions (reviewed in Hilliker *et al.*, 1980) are also located in these regions. It is now of importance to determine if these genetic functions are encoded by unique sequences functioning amongst non-transcribed repeated sequences.

Despite extensive knowledge of the DNA sequence organization of heterochromatin, very little is known about the higher order structure of chromatin that may also be responsible for the unusual properties of the heterochromatin. The phasing of nucleosomes may differ between actively transcribed, and nontranscribed, chromosomal regions (Zachau and Igo-Kemenes, 1981; Weisbrod, 1982) and certain chromosomal proteins are preferentially located in nucleosomes from each type of region. For instance, ubiquitin is found in nucleosomes isolated from transcriptionally active chromatin (Levinger and Varshavsky, 1982(a)), whereas a protein termed D1 (Alfamgeme *et al.*, 1980), is associated with nucleosomes containing satellites 1 and 111 (Levinger and Varshavsky, 1982(b)), which contain 100% and 69% A-T, respectively. In addition to binding to satellite sequences, the D1 protein may also be associated with other AT-rich tandemly repeated regions. Examination of the distribution of chromosomal proteins on polytene chromosomes using immunofluorescence has identified three further proteins associated specifically with heterochromatic sequences. One such protein recognises the nucleolus organizer (Weideli *et al.*, 1978), a second, isolated

from *D. melanogaster*, hybridizes to  $\alpha$ -heterochromatin more strongly than  $\beta$ -heterochromatin in *D. virilis* (Will and Bautz, 1980) and the other is specifically associated with sequences from satellite III (Hsieh and Brutlag, 1979(a)). Analysis of the sequence recognised by satellite DNA-binding proteins has revealed a highly symmetric region (Brutlag, 1980) similar to those sites recognised by prokaryotic DNA binding proteins. Despite much speculation, the function of these proteins is unknown, although their analysis may provide further insights into the distinctive properties of heterochromatin.

### 1.1.3 Intercalary heterochromatin

Sites of heterochromatin within the euchromatin, termed 'intercalary heterochromatin', were initially identified on the basis of three criteria: high frequency of breakage, ectopic pairing and differential response to alkali-urea treatment (Hannah, 1951). Cytologically, some intercalary heterochromatic regions appear as capsules or bulbs and weak spots, suggesting that they contain repeated sequences. More recently, sites of intercalary heterochromatin of polytene chromosomes have been shown to replicate late and contain repeated elements, as determined by *in situ* hybridization to cRNA and total RNA (Zhimulev *et al.*, 1982). Furthermore, Zhimulev *et al.* (1982) has demonstrated a significant correlation in the manifestation of the described properties of intercalary heterochromatin, although not every feature is equally expressed at all locations. The analysis also suggested that the characteristics of intercalary heterochromatin are dependent on its DNA sequences.

The evidence suggests that intercalary heterochromatin consists of clusters of tandemly repeating units. Both the histone and 5S RNA genes are organized into tandem arrays (Lifton *et al.*, 1977; Procunier and Dunn, 1978)



and are located at regions of intercalary heterochromatin (Ananiev *et al.*, 1978). In addition, a component of satellite IV, together with sequences transcribed at relatively high levels in cultured cells (Gvozdev *et al.*, 1980), are present at genetically defined sites of intercalary heterochromatin. As transcripts homologous to at least some mobile elements are present at high levels in the polyadenylated RNA fraction of cultured cells (Schwartz *et al.*, 1982), the chromosomal locations of mobile element families were tested to assess if their locations were correlated with sites of intercalary heterochromatin. The distribution of two elements, *mdg-1* and *mdg-3*, was found to be non-random and could be correlated with intercalary heterochromatic sites (Ananiev *et al.*, 1978; Belyaeva *et al.*, 1984). However, the data on the chromosomal location of the *copia*, 412 and 297 families were contradictory (Young and Schwartz, 1980; Strobel, 1982; Montgomery and Langley, 1983) although DNA sequencing has revealed that some mobile elements integrate into AT-rich regions, which are characteristic of satellites (for example, Ikenaga and Saigo, 1982; Snyder *et al.*, 1982; McGinnis *et al.*, 1983(a); Inouye *et al.*, 1984).

Lifschytz (1983) has compared the degree of replication in diploid and polytene tissues of sequences representing the unique, middle repetitive from both euchromatin and heterochromatin, and tandemly repeated classes. The data indicate that under-replication is a function of tandem arrays, further supporting the proposal that the properties of intercalary heterochromatin are due to tandem arrays of repeated sequences.

## 1.2 Chromosome Position and Gene Function

### 1.2.1 Position effect and variegation

Genetic studies have indicated that the chromosomal environment of a particular gene is important to its functioning, since movement of a gene to

different positions in the genome can alter its expression. Heterochromatic position effect is defined as the alteration in the expression of a euchromatic gene, which has been placed adjacent to, or inserted into, heterochromatin (reviewed in Baker, 1968; Spofford, 1976). In order to obtain position effects, there are two stringent requirements: a break must have occurred within heterochromatin and the chromosomal rearrangement must have placed the affected gene close to the breakpoint i.e. in *cis* rather than *trans* configuration. Where the altered gene function involves cell autonomous phenotypes, the position effect is visualized as a variegated phenotype (e.g. white-mottled), whereas if the position effect involves non-autonomous cell functions (e.g. the amylase locus), it is seen as reduced expression per individual. Variegation of genes normally located in a heterochromatic region occurs if they are placed in a distal euchromatic region or into disrupted heterochromatin. The euchromatic and heterochromatic breakpoints of three variegating *white* mutants have been examined by Tartof *et al.* (1984). The euchromatic breakpoints were found to be located approximately 25kbp downstream from the *white* structural gene, thus confirming the hypothesis, proposed from genetic data, that position effect does not result from gene mutation. The three breakpoints were clustered within 1kbp and were flanked by sequences homologous to the 1.688g/cc satellite which suggested that chromosome breakage may have occurred after ectopic pairing of this region with satellite sequences in the centromeric heterochromatin. The heterochromatic breakpoints were located in middle repetitive sequences: two occurred in type 1 insertion sequences of ribosomal RNA genes and the other in a sequence with characteristics of mobile elements.

Characteristically, *cis*-inactivation of DNA sequences by adjoining heterochromatin extends beyond the limits of the adjacent locus (Spofford, 1976). Variegation is most frequent for the euchromatic locus closest to the

breakpoint and diminishes with distance from it.<sup>\*</sup> The developmental time at which the variegation event occurs is variable (Baker, 1968). In some cases, the mosaicism is fine grained, indicating the decision to inactivate the relocated locus was made late in development, whereas in other instances, the inactive patches are clonal descendant of cells in which the decision was made early in development.

In order to explain the alteration in phenotype on changing the chromosomal position of a gene, two hypotheses have been proposed. The first suggests that somatic gene loss occurs in some cells if genes placed adjacent to centromeric heterochromatin, are drawn into a domain of under-replication (Schultz, 1936). The alternative proposal suggests that the gene is inactivated in a portion of the tissue only (Baker 1968). The two proposals have been tested using molecular techniques. Relocation of the *rosy* (Rushlow *et al.*, 1984) and 87C heat shock (Henikoff, 1981) gene sequences near to chromocentric heterochromatin resulted in altered gene expression. These genes were not under-represented at their new chromosomal locations, thus refuting the somatic gene loss model. For the *rosy* gene, it has been demonstrated directly that the number of homologous transcripts was greatly reduced.

It is possible that position effect results from the disruption of the initiation of transcription, due to a lack of accessibility of the required transcription factors. Several further observations implicate chromatin structure in controlling position effect. Specific non-histone proteins are associated with satellite and heterochromatic sequences (Section 1.1.1) and deletions of the histone genes (Moore *et al.*, 1983), or chemical modification of histones (Mottus *et al.*, 1980), enhance the activity of the variegating gene. Henikoff (1979) has identified and localized enhancers of variegated gene function and suggested that they may represent genes coding for

\* Several environmental factors, for example temperature (Hartmann-Goldstein, I., *Genetical Research* 10:143-159, 1967), influence the extent of variegation.

chromosomal proteins. The chromatin structure may, therefore, differ between euchromatin and heterochromatin, due to the binding of different proteins to each chromosomal state.

Tartof *et al.* (1984) have proposed a model to account for heterochromatic domains. Two sites were defined, one at which the heterochromatic state is initiated (*i*), and a second, at which it is subsequently terminated (*t*). Although a number of mechanisms have been described for the initiation of heterochromatinization (Gartler and Riggs, 1983), Tartof *et al.* (1984) have proposed that heterochromatin forms when *i* sites bind proteins that interact with each other in a cooperative manner, generating a compacted chromosome domain that extends until a termination site is reached. The model can be directly applied to a euchromatin-heterochromatin transition zone, where the *i* and *t* sites may represent the boundaries of the heterochromatic blocks adjacent to euchromatin. Therefore, only the heterochromatic domains would have a compacted structure.

#### 1.2.2 Transformation studies

In contrast to analysing euchromatin-heterochromatin chromosomal rearrangements, the effect of chromosome position on gene function can be studied more directly by inserting genes into different chromosomal locations. This can be accomplished using the P-element mediated transformation system, in which *Drosophila* embryos are injected with DNA of a P element vector carrying defined sequences (Spradling and Rubin, 1982; Rubin and Spradling, 1982). The sequences are efficiently and selectively transposed from extra-chromosomal DNA into *Drosophila* germ-line chromosomes. The integrated sequence is not rearranged and is stably inherited.

Reintroduction of a number of genes into the *Drosophila* genome at a variety of chromosomal locations, has demonstrated that only a small amount of

flanking DNA is required for correct tissue and temporal specificity, and for synthesis of approximately wildtype levels of the gene product (for example, Spradling and Rubin, 1983; Goldberg *et al.*, 1983; Scholnick *et al.*, 1983; Richards *et al.*, 1983; Bargiello *et al.*, 1984). There have been only four reported instances of the insertion of a gene into heterochromatin, and in each case gene activity was disrupted. In two instances, the insertion occurred into chromocentric heterochromatin, whereas in the other two examples the insertion occurred into telomeric sequences (Spradling and Rubin, 1983; Hazelrigg *et al.*, 1984; Gehring *et al.*, 1984). However, there have also been a number of examples of the disruption of only some aspects of a gene's function after reintroduction of the gene. For example, transformants containing the alcohol dehydrogenase (*Adh*) gene inserted into the X chromosome did not undergo dosage compensation, while two other *Adh* transformants showed reduced expression in either the larval or adult stage (Goldberg *et al.*, 1984); many *rosy* transformants displayed reduced xanthine dehydrogenase levels (Spradling and Rubin, 1983); the majority of reinserted *white* genes did not respond to *zeste* in a wildtype manner (Gehring *et al.*, 1984; Hazelrigg *et al.*, 1984) and the amplification of inserted chorion genes in follicle cells was variable (de Cicco and Spradling, 1984). In most of these examples there was no clear correlation between the disruption of gene activity and chromosome position, except in the case of the *rosy* transformants for which there was an association between reduced gene activity and insertion at sites of intercalary heterochromatin.

Relatively few P-element mediated transformants have been recovered in which the inserted gene is non-functional, probably because most transforming protocols identify integration on the basis of gene activity. However, chorion gene transformants were not selected in this manner, with the result that 44% of transformants did not show wildtype amplification of the inserted



chorion genes in the follicle cells. In this experiment, low levels of gene activity were correlated with the presence of heterochromatin at the site of insertion in 59% of the transformants.

P-element mediated transformation studies have allowed two conclusions to be drawn concerning the effect of chromosomal position on gene function. Firstly, it is clear that, as the inserted sequence was not expressed in a wildtype manner at all locations, regulatory features of some chromosomal positions can over-ride the *cis*-acting controlling sequences carried with the gene. Secondly, although gene expression was correlated with the chromatin state into which the gene inserted in the majority of transformants, at some sites this factor alone is insufficient to influence the level of gene expression.

### 1.3 Molecular Analysis of the Structural Organization and Regulation of Genes from *D.melanogaster*

#### 1.3.1 Structural organization of cloned genes

Molecular cloning techniques were initially used to isolate genes whose RNA products were abundant in at least some cell types. When the genes encoding well-defined, highly abundant protein products are considered as a group, the size of the genes, as determined by the distribution of coding sequences, and the size of transcripts, are both relatively small (Table 1.1). This reflects the low number of introns (0 to 3) and the small intron sizes. The genes which represent exceptions to this organization at present are the myosin heavy- and light-chain genes, a collagen-like gene and the genes coding for GAR transformylase and tropomyosin 1. This group of genes appears to be organized into several discrete classes. The simplest class is the single copy gene, as exemplified by *Adh*. However, if a single copy gene generates a product required either in different tissues or alternative life

stages, multiple transcripts may be produced from the same DNA sequence. A second class, termed 'clustered multicopy' (Spradling and Rubin, 1981), is characterized by the clustering, within a short stretch of DNA, of genes coding for functionally related proteins. Within this category fall the genes coding for the larval cuticle proteins, histones, tropomyosins, four small heat-shock proteins, three salivary gland glue proteins genes and two clusters of chorion genes. Alternatively, the genes encoding functionally related products may be dispersed around the genome, as illustrated by the actin, tubulin and larval serum protein genes. As some genes are organized as dispersed clusters, they share the properties of the latter two classes (for example, yolk protein and *hsp 70* genes).

The diversity of organization is clearly demonstrated by genes coding for four members of the contractile protein family: the actins, myosin heavy and light chains, and the tropomyosins, each of which falls into a different organizational class. It is not clear why some genes are organized into small clusters while others are dispersed, although clusters of functionally related genes are often under coordinate regulation. Several examples have been described in which related members of a gene cluster are transcribed in opposite directions, with the sequence in between proposed to be important in the coordinate regulation. Such examples include histone, larval cuticle protein and yolk protein genes (Lifton, 1977; Riddell, 1981; Snyder *et al.*, 1982). It has been demonstrated that the sequences between the yolk protein genes, *yp1* and *yp2*, regulate the tissue specific transcription of both genes. Separate elements that regulate expression in the fat body and the ovaries have been identified (Garabedian *et al.*, 1985). However, there are examples of dispersed genes being coordinately controlled. For example, the heat shock proteins and subunits of larval serum protein 1 (Holmgren *et al.*, 1979; Smith *et al.*, 1981; Ingolia and Craig, 1982).

More recently, techniques have been devised to allow the isolation of genes whose products are either unknown, or present in low abundance. These techniques include chromosome walking (Bender *et al.*, 1983(a)), transposon tagging (Spradling and Rubin, 1982) and micro-cloning (Pirrotta *et al.*, 1983). Genes classified as 'complex' in terms of their genetic organization and developmental phenotypes, have now been isolated, and in all cases the size of the gene is large compared to that of genes encoding the highly abundant proteins (Table 1.1). This property is especially striking for the *Notch* and *scute* loci, which extend over 40 and 60kbp respectively. Another class of genes, the homeotic genes, specify segmental identity and two genes from this class have been cloned, namely, *bithorax* and *Antennapedia*. Both are characterized by their large size, with the *Antennapedia* gene, extending over 100kbp, being the largest *Drosophila* gene yet reported. Both the complex and homeotic loci encode multiple transcripts and each has a number of introns, some of which may be large. Transcription from the *Notch* locus is not well defined as yet, although this gene appears to encode a major transcript of 10.5kbp and a number of minor species. Likewise, the *white* locus produces two transcripts of 2.6 and 2.0kbp, as well as minor species. The homeotic genes each encode three transcripts that are relatively small compared to the size of the gene, and each transcript contains large multiple introns (Table 1.1). In contrast to the complex and homeotic loci, those genes required in early embryogenesis for the formation of the segmental pattern appear to be variable in terms of their organization. Although *engrailed* is sometimes considered to be a homeotic gene, it appears to belong to the 'pair-rule' class of segmentation genes (Nusslein-Volhard and Wieschaus, 1980). The *fushi-tarazu* gene also belongs to the 'pair-rule' class, whereas *Kruppel* belongs to the 'gap' class of segmentation genes. The *dorsal* gene is the only maternal effect gene characterized to date and is required for the



establishment and maintenance of the anterior-posterior and dorsal-ventral polarity. The latter four genes vary in size from 2.0 to 40kbp and may produce a single or multiple transcripts (Table 1.1).

Another category of genes to consider are those broadly defined as neurobiological. These include *dunce* and dopa decarboxylase, both of which affect learning (Dudai *et al.*, 1976; Tempel *et al.*, 1984) and *period* which is one of the loci controlling biological rhythms (Hall, 1984). Two loci of the complex class can also be considered in this category, as *scute* is involved in the differentiation of innervated elements in both the peripheral and central nervous systems (Garcia-Bellido and Santamaria, 1978), and *Notch* may determine differentiation of ectodermal cells into neuroblasts (Lehmann *et al.*, 1983). The most outstanding feature of this class of genes is their overall complexity. Each gene, except dopa decarboxylase, is large, extending over 20 to 60kbp. All genes produce multiple transcripts, although in the case of dopa decarboxylase and *period*, there is some uncertainty as to which transcripts are functional.

### 1.3.2 Regulation of gene expression

Genetic studies have suggested that *cis*-acting regulatory sequences are associated with some structural genes. This is well exemplified by the *white* and *rosy* loci, for which fine structure mapping studies have demonstrated that mutations fall into either a structural or regulatory domain (Judd, 1976; Chovnick *et al.*, 1977; McCarron *et al.*, 1979). The ability to clone genes and determine their structural organization has enabled the genetic predictions to be tested at the molecular level. Molecular analysis of the *white* gene has demonstrated that several mutations map to either the 5' region of the transcription start site ( $w^{sp}$ ,  $w^{DZL}$ ) or the 5' untranslated region ( $w^e$ ) (O'Hare *et al.*, 1984). Alterations in either region disrupts transcription

(Levis *et al.*, 1984) and P-element mediated transformation of the *white* gene has confirmed the requirement of these sequences for approximately wildtype levels of transcription. Similarly, analysis of variants which map to the regulatory region of the *rosy* locus has identified sequences which regulate transcription (Clark *et al.*, 1984).

Several alternative experimental approaches have been utilized to examine the role of *cis*-acting sequences in regulating gene expression, and these are particularly applicable to genes for which mutations have not been characterized. The first of these techniques generates deletions of the 5' regions flanking the transcription start site, and gene activity is then assayed using the P-element mediated transformation system. This method has allowed the identification of *cis*-acting regulatory sequences which control the level of transcription of the rRNA (Kohorn and Rae, 1983), dopa decarboxylase (Hirsh, 1984) and *Sgs-3* genes (Baurouis and Richards, 1985) and those which are responsible for the developmental amplification of the chorion genes (de Cicco and Spradling, 1984). For those cases in which the molecular organization of a gene is well defined, detailed analyses of sequences responsible for differential expression have been possible. For example, the alcohol dehydrogenase gene contains alternative promoters for use at the larval and adult stages (Benyajati *et al.*, 1983). Deletion analysis has located the sequences required for larval expression to a region 1.4kbp 5' to the adult promoter, although the adult promoter is the more distal, whereas the adult specific sequences lay 60bp 5' to the adult promoter. In addition, those sequences required for adult expression have enhancer-like properties, as when fused to other promoters (e.g. alcohol dehydrogenase larval or dopa decarboxylase promoters) gene expression is increased (Posakony, 1984).

A second approach to identifying *cis*-acting regulatory sequences utilizes the observation that chromatin, when being actively transcribed, is

differentially sensitive to DNA'ase 1. Extensive analysis of DNA'ase 1 hypersensitive sites surrounding the salivary gland glue protein gene, *Sgs-4*, has demonstrated the presence of five hypersensitive sites in the 5' region in chromatin prepared from salivary glands, with those at -405bp and -330bp being the most prominent (Shermoen and Beckendoff, 1982). Analysis of strains displaying variation in *Sgs-4* production has revealed a strong correlation between the presence of DNA'ase 1 hypersensitive sites in chromatin and gene expression (McGinnis *et al.*, 1983(a); McGinnis *et al.*, 1983(b)).

DNA'ase 1 hypersensitive sites have also been found associated with the 5' flanking regions of the heat shock protein genes, *hsp70* and *hsp82* (Wu, 1980). One of these sites is located at, or near, the TATA box. The role of the other sites has been examined by generating a series of deletions in the 5' region and testing for gene activity in either a heterologous system (e.g. Corces *et al.*, 1981; Pelham, 1982; Mirault *et al.*, 1982) or using the P-element mediated transformation system (Cohen and Meselson, 1984; Dudler and Travers, 1984). A correlation was demonstrated between hypersensitive sites and sequences necessary to elicit the heat shock response. The role of these sequences in regulating the expression of the *hsp82* gene, has been further refined by the identification of proteins bound at, or near, the hypersensitive site. A protein was bound at the TATA box both before, and after, heat shock whereas a second protein, termed HAP, was bound to the upstream sites only after heat shock (Wu, 1984(a); 1984(b)). The protein binding sequence shares homology with the 5' region of a subset of related heat shock proteins.

Despite the identification of *cis*-acting regulatory sequences, it is still not clear how genes are stably repressed under certain conditions and activated in response to *trans*-acting factors under others. Recent evidence suggests that gene expression may be modulated by interactions between

different RNA products , between RNA and DNA and between DNA and proteins. Very little is known of the role of RNA products in regulating gene expression. However, it has been suggested that in some prokaryotic systems, anti-sense RNA binds with the functional transcript rendering it inactive (Mizundo *et al.*, 1984; Kumar *et al.*, 1985; Keimar and Novick, 1985). Although there is no evidence to date that this form of regulation operates in eukaryotes, it has proven possible to block translation of certain mRNAs by injection of anti-sense RNA into either vertebrate cells (Izant and Weintraub, 1984; Melton, 1985) or *Drosophila* (Rosenberg *et al.*, 1985). There is also evidence from the variant surface antigen genes of *Trypanosoma brucei* to suggest that mini-exons, located at considerable distances from the remaining exons, are transcribed separately. The RNA products of these mini-exons may control transcription of the longer segment (Campbell *et al.*, 1984).

The role of proteins in regulating eukaryotic gene expression has recently become clearer. Studies of *Xenopus* 5S ribosomal genes have led Brown (1984) to propose that a gene can be transcriptionally active only in the presence of a transcription complex consisting of a number of proteins interacting strongly with each other and the control region of the gene. The likelihood of a stable transcription complex forming would depend on the relative concentration of the required factors and repressor molecules, and the binding affinities of each type of molecule for the gene. In this condition, the gene is 'determined', or committed to expression under defined conditions and the determined state is inherited through cell divisions. The gene would be activated under the influence of additional cellular modulating factors, such as the RNA polymerase and effector molecules including hormones or hormone receptors.

Transcription factors have now been identified for other genes transcribed by RNA polymerase III (Lassar *et al.*, 1983), as well as for

several transcribed by RNA polymerase II (Davison *et al.*, 1983). Two factors that may form part of transcription complexes have been identified in *Drosophila*: one is specific for the promotor region of the histone H3, H4 and actin 5C genes (Parker and Topol, 1984(a)), whereas the other is found only in heat-shocked cells and appears to bind to a DNA sequence upstream from the *hsp* 70 promotor (Parker and Topol, 1984(b)).

Proteins mediating *trans* regulation have been described for the classical prokaryotic systems, such as the *cl* and *cro* repressors of bacteriophage and the *lac* repressor and catabolite gene activator of *E.coli* (Gilbert *et al.*, 1976; Takeda *et al.*, 1983). More recently, *trans*-acting regulators of gene expression have been described in eukaryotes. For example, human T-cell leukemia viruses (HTLV-I and -II) produce a factor which stimulates transcription from the viral promotor in the long terminal repeat (Sodroski *et al.*, 1984). In addition, expression of the glue protein gene, *Sgs-3*, of *D. melanogaster* requires the product of a gene located at polytene chromosome band 2B5 (Crowley *et al.*, 1984). However, it has not yet proven possible to correlate those genes defined genetically as having a regulatory function, with a product that acts in *trans* to influence gene expression.

A *trans*-acting regulatory function has been suggested for a sequence recently isolated from *Drosophila*. The sequence, termed the 'homeo-box', has been localized to the 3' end of mature transcripts from the related homeotic genes (McGinnis *et al.*, 1984; Scott and Weiner, 1984; Gehring, 1985), and other developmental genes including *fushi-tarazu* (Laughton and Scott, 1984) and *engrailed* (Fjose *et al.*, 1985; Poole *et al.* 1985). This sequence has been conserved across all vertebrate and many of the invertebrate phyla surveyed to date (reviewed in Gehring, 1985). The open reading frame of the sequence would encode a product with properties characteristic of DNA-binding proteins. It has therefore been suggested that the homeo-box codes for a



conserved DNA-binding moiety which is tagged onto a diverse set of sequences to regulate gene expression in *trans*.

The studies described above suggest that protein-DNA interactions modulate the efficiency of transcription initiation by specifically altering DNA conformation and/or chromatin structure, thereby facilitating the activity of nearby promoters. This proposal has been supported by analysis of several gene systems in other organisms, for example, mouse mammary tumour virus (Zaret and Yamamoto, 1984). It also appears that the DNA conformation in the vicinity of regulatory sequences may be altered in an actively transcribed gene. This is illustrated by the identification of sequences with the potential to form Z-DNA in the regulatory region of the *Drosophila hsp 70* gene (Nussinov and Lennon, 1984) and the simian virus 40 (SV40) enhancer element (Nordheim and Rich, 1983). Hypersensitivity to S1 Nuclease suggests the formation of cruciform structures and such sites have been identified in the regulatory regions of genes such as  $\epsilon$ -globin (Allan *et al.*, 1984), rat preproinsulin 11 and SV40 origin (Evans *et al.*, 1984).\*

Analysis of eukaryotic gene systems, in organisms apart from *Drosophila*, have identified three types of sequences involved in regulating transcription. The first type, termed the TATA box, positions the transcription start site 30bp further downstream of the TATA box. A second element, termed the CCATT box, is located approximately 40 to 110bp 5' to the transcription start site and is thought to be important in determining the level of transcription. The third type of regulatory sequence is the enhancer element, which may play a role in determining tissue-specific modulation. Apart from the TATA box, very few elements have been correlated with defined functions in regulating the expression of *Drosophila* genes. Furthermore, although regulatory sequences have also been identified 3' to the transcription start site in some higher eukaryotes (for example, Charney *et*

\* However, DNA regions undergoing a B to Z transition, and other specific DNA structures, are also sensitive to this enzyme (Wells, R.D. *et al.*, Cold Spring Harbor Symp. Quant. Biol. 47:77-84, 1982).

*al.*, 1984; Wright *et al.*, 1984), such sequences have not yet been located in *Drosophila*. In addition, the involvement of *trans*-acting factors, particularly hormones, in the initiation of gene expression has been revealed in studies of higher eukaryotes (for example, Karin *et al.*, 1984; Moore *et al.*, 1985). Although it is known that ecdysteroid hormones stimulate the expression of some genes in *Drosophila* (reviewed in Yund and Germeraad, 1980), the mechanism by which this occurs is poorly understood. However, the identification of ecdysone inducible proteins and the localization of the genes encoding these proteins (Savakis *et al.*, 1984), together with the identification of genes whose products affect hormonally controlled genes (Crowley *et al.*, 1984), suggests that this regulatory mechanism may soon be elucidated.

#### 1.4 The Proximal Region of the X-Chromosome: A Molecular and Genetic Analysis

The proximal region of the X-chromosome of *D. melanogaster*, and in particular divisions 19 and 20, was chosen to study the organization and regulation of genes in a euchromatin-heterochromatin transition zone. In this region, the polytene chromosome organization is changing from cytologically distinguishable euchromatin to heterochromatin and the two divisions lie adjacent to the centromeric heterochromatin, although it has not been possible to determine the location of heterochromatin, as defined on a mitotic chromosome, on the polytene chromosome. Cytogenetic analyses have shown that gene sequences extend through divisions 19 and 20 (Lifschytz, 1971; Schalet and Lefevre, 1973; 1976; Lefevre, 1981; Kramers *et al.*, 1983) and although Schalet and Lefevre (1976) have proposed that the euchromatin-heterochromatin junction in the mitotic chromosome lies proximal to the *suppressor of forked* locus, the nature of the junction is unknown (Lifschytz, 1978). The sites of intercalary heterochromatin dispersed within the euchromatin of the polytene

chromosome have been inferred from an analysis of the distribution of radiation induced chromosomal rearrangements along the X chromosome (Figure 1.2). Subdivisions of the polytene chromosome constituting 'hotspots' for chromosome breakage are probable sites of intercalary heterochromatin.

Divisions 19 and 20 have been subjected to extensive genetic analysis (Lifschytz and Falk, 1968; 1969; Lifschytz, 1971; 1978; Schalet and Lefevre, 1973; 1976; Lefevre, 1981; Kramers *et al.*, 1983), and, consequently the region is considered to have been saturated with contiguous lethal complementation groups. Chromosomal rearrangements have also been defined and this wealth of genetic information enables genetic manipulations of the region to be undertaken with comparative ease.

Divisions 19 and 20 are also of interest because of the clusters of visible mutations that have been localized to the region. Of the mutations which have been mapped to complementation groups, the phenotype of a number were found to be duplicated within the region (Figure 1.3). For instance, the mutant phenotypes of *melanized/melanized-like*, *little fly/little fly-like*, *uncoordinated/uncoordinated-like*, *outheld/wings apart* are all similar. Complementation groups exhibiting similar mutant phenotypes may have arisen as a result of gene duplication.

In addition to those loci indicated on Figure 1.3, recombination mapping studies have localized clusters of genes to divisions 19 and 20 that may affect developmental or neurobiological functions (Table 1.2). At least four genes have been identified which may affect early embryonic developmental decisions, including *tumourous head I*, which influences the anterior-posterior gradient; *runt*, involved in the formation of segments; *folded gastrulation*, associated with morphological movements at gastrulation and *extra organs*, which may influence cell determination in some lineages. However, the most abundant gene cluster is that affecting neurobiological and muscle functions



(reviewed in Hall, 1982) and there appears to be a disproportionate number of this class of genes in regions of divisions 19 and 20. The genes can be classified according to functions associated with the visual system, flight and thoracic muscles and physiological and motor coordination. The latter category is the largest, and it is into this class that the *uncoordinated* gene, analysed in this thesis, belongs.

Mutations at the *uncoordinated* locus are semi-lethal, and flies carrying the mutation are unable to walk because of lack of leg coordination, and if they eclose, die soon after. Consequently, *uncoordinated* belongs to that class of gene which may be responsible for determining muscle and/or neuromuscular functions, and as such, may be of importance in determining motor coordination. The analysis of such loci may lead to an understanding of the function, development and integration of sensory, neural and muscular systems.

The work presented in this thesis describes the isolation of the *uncoordinated* gene and its flanking sequences by the technique of chromosome walking. The molecular organization of the gene and surrounding sequences in a euchromatin-heterochromatin transition zone was analysed and compared to that of other genes isolated from divisions 19 and 20, and to genes which have been characterized from more conventional euchromatic regions. Some insight into the manner of regulating tissue and temporal specificity of gene expression was attained by identifying the transcriptional unit of the *uncoordinated* gene. The data was supported by the identification of the molecular lesion in alleles mutated at the *uncoordinated* gene, and an examination of the relationship between the structural defect and the transcriptional properties of the mutated gene.

Table 1.1

Genes cloned from *D. melanogaster*

| Gene                       | Size of Gene<br>(kbp) | Number of<br>Transcripts | Size of<br>Transcripts | Number of<br>Introns | Chromosomal<br>Location | Reference  |
|----------------------------|-----------------------|--------------------------|------------------------|----------------------|-------------------------|--|
| a) Highly Abundant Protein |                       |                          |                        |                      |                         |  |
| i) single copy             |                       |                          |                        |                      |                         |  |
| Alcohol dehydrogenase      | 2.0                   | 2                        | 1.07<br>1.01           | 3<br>2               | 35B2-3                  | Goldberg, 1980<br>Benyajati <i>et al.</i> , 1983                   |
| GAR transformylase         | 6.9                   | 2                        | 4.7<br>1.7             | 5<br>2               | N.D.                    | Henikoff <i>et al.</i> , 1983                                      |
| Myosin heavy chain         | 19                    | 3                        | 8.6<br>8.0<br>7.2      | 9<br>8<br>7          | 36B                     | Rozek and Davidson, 1983   |
| Myosin light chain         | 3.5                   | 2                        | 1.3<br>0.9             | 5<br>4               | 98B                     | Falkenthal <i>et al.</i> , 1984<br>Falkenthal <i>et al.</i> , 1985 |
| rosy                       | 5.0                   | 1                        | 4.7                    | ?                    | 87D10-12                | Clark <i>et al.</i> , 1984   |
| ii) dispersed multicopy    |                       |                          |                        |                      |                         |  |
| Actin 1                    | 1.73                  | 1                        | 1.7                    | 1                    | 88F                     | Fyrberg <i>et al.</i> , 1980                                       |
| Actin 2                    | 3.3                   | 3                        | 2.2<br>1.95<br>1.7     | at least 1           | 5C                      | Tobin <i>et al.</i> , 1980   |
| Actin 3                    | 1.7                   | 1                        | 1.7                    | 0                    | 42A                     | Fyrberg <i>et al.</i> , 1981                                       |
| Actin 4                    | 2.2                   | 1                        | 1.75                   | 1                    | 57A                     | Fyrberg <i>et al.</i> , 1983                                       |
| Actin 5                    | 1.6                   | 1                        | 1.6                    | 0                    | 87E                     |  |
| Actin 6                    | 1.9                   | 1                        | 1.6                    | 1                    | 79B                     |  |
| Collagen-like 1            | N.D.                  | 1                        | N.D.                   | N.D.                 | 19EF-20AB               | Natzle <i>et al.</i> , 1982  |
| Collagen-like 2            | N.D.                  | 1                        | 6.4                    | N.D.                 | 25C                     | Monson <i>et al.</i> , 1982  |
| $\alpha$ -tubulin-1        | N.D.                  | 1                        | 1.8                    | N.D.                 | 84B3-C8                 | Kalfayan and Wensink, 1982   |
| $\alpha$ -tubulin-2        | N.D.                  | 1                        | 1.65                   | N.D.                 | 85E6-F5                 | Natzle and McCarthy, 1984  |
| $\alpha$ -tubulin-3        | N.D.                  | 1                        | 2.0                    | N.D.                 | 84D5-8                  |  |

|                       |      |      |      |      |           |                                |
|-----------------------|------|------|------|------|-----------|--------------------------------|
| $\alpha$ -tubulin-4   | N.D. | 4    | 1.95 | N.D. | 67C4-6    |                                |
|                       |      |      | 1.9  | N.D. |           |                                |
|                       |      |      | 1.7  | N.D. |           |                                |
|                       |      |      | 1.5  | N.D. |           |                                |
| $\beta$ -tubulin-1    | 1.8  | 1    | 1.8  | 0    | 56C       | Natzle and McCarthy, 1984      |
| $\beta$ -tubulin-2    | 2.2  | 1    | 2.2  | 0    | 85E       | Pialoan <i>et al.</i> , 1984   |
| $\beta$ -tubulin-3    | 2.5  | 1    | 2.5  | 0    | 60C       |                                |
| $\beta$ -tubulin-4    | 1.8  | 1    | 1.8  | 0    | 97EF      |                                |
| LSP-1a                | 2.85 | 1    | 2.85 | 0    | 11A7-B9   | Smith <i>et al.</i> , 1981     |
| LSP-1 $\beta$         | 2.85 | 1    | 2.85 | 0    | 21D2-22A1 |                                |
| LSP-1y                | 2.85 | 1    | 2.85 | 0    | 61A1-A6   |                                |
| LSP-2                 | N.D. | 1    | 1.8  | N.D. | 68EF      |                                |
| Yolk Protein 1        | N.D. | 1    | 1.6  | 1    | 8F-9A     | Barnett <i>et al.</i> , 1980   |
| Yolk Protein 2        | N.D. | 1    | 1.6  | 1    | 8F-9A     | Riddell <i>et al.</i> , 1981   |
| Yolk Protein 3        | N.D. | 1    | 1.6  | 1    | 12B-C     |                                |
| Vitelline Membrane-1  | N.D. | 1    | 0.65 | N.D. | 34C       | Mindrinos <i>et al.</i> , 1985 |
| Vitelline Membrane-2  | N.D. | 1    | 0.75 | N.D. | 26A       |                                |
| Vitelline Membrane-3  | N.D. | N.D. | N.D. | N.D. | 32EF      |                                |
| Vitelline Membrane-4  | N.D. | N.D. | N.D. | N.D. | 26A       |                                |
| Ribosomal protein 7/8 | N.D. | 3    | 1.8  | N.D. | 5D        | Burns <i>et al.</i> , 1984     |
|                       |      |      | 1.6  | N.D. | 5D        |                                |
|                       |      |      | 1.2  | N.D. | 5D        |                                |
| S18                   | N.D. | 1    | 0.66 | N.D. | 15B       |                                |
| L12                   | N.D. | 1    | 1.05 | N.D. | 62E       |                                |
| r49                   | N.D. | 1    | 0.65 | N.D. | 99D       | Vaslet <i>et al.</i> , 1980    |
| Ecdysone-induced      | N.D. | 1    | 2.7  | N.D. | 74EF      | Moritz <i>et al.</i> , 1984    |

iii) clustered multicopy

|                   |      |      |           |      |            |                                    |
|-------------------|------|------|-----------|------|------------|------------------------------------|
| Chorion s38       | N.D. | 1    | 1.4       | N.D. | 7F1-2      | Spradling <i>et al.</i> , 1980     |
| Chorion s36       | N.D. | 1    | 1.2       | N.D. | 7F1-2      | Spradling, 1981                    |
| Chorion s16       | 0.70 | 1    | 0.65      | 1    | 66D11-15   | Griffin-Shea <i>et al.</i> , 1980  |
| Chorion s19       | 0.80 | 1    | 0.75      | 1    | 66D11-15   | Griffin-Shea <i>et al.</i> , 1982  |
| Chorion s15       | 0.77 | 1    | 0.65      | 1    | 66D11-15   |                                    |
| Chorion s18       | 0.85 | 1    | 0.75      | 1    | 66D11-15   |                                    |
| Cuticle protein-1 | 0.56 | 1    | 0.5       | 1    | 44D        | Snyder <i>et al.</i> , 1981        |
| Cuticle protein-2 | 0.56 | 1    | 0.5       | 1    | 44D        | Snyder <i>et al.</i> , 1982        |
| Cuticle protein-3 | 0.56 | 1    | 0.5       | 1    | 44D        |                                    |
| Cuticle protein-4 | 0.56 | 1    | 0.5       | 1    | 44D        |                                    |
| Cuticle protein-5 | N.D. | 1    | 1.0 + 1.5 | N.D. | 44D        |                                    |
| Sgs-1             | N.D. | N.D. | N.D.      | N.D. | 25B        | Crowley <i>et al.</i> , 1983       |
| Sgs-3             | 1.17 | 1    | 1.1       | N.D. | 68C        |                                    |
| Sgs-4             | 0.95 | 1    | 0.95      | 0    | 3C11-12    | Muskavitch and Hogness, 1980       |
| Sgs-5             | N.D. | 1    | 0.75      | N.D. | 90B3-8     | Guild, 1984; Guild and Shore, 1984 |
| Sgs-6             | N.D. | N.D. | N.D.      | N.D. | 71C-F      | Velissariou and Ashburner, 1981    |
| Sgs-7             | 0.42 | 1    | 0.36      | N.D. | 68C        | Crowley <i>et al.</i> , 1983       |
| Sgs-8             | 0.39 | 1    | 0.32      | N.D. | 68C        |                                    |
| EIP 28            | N.D. | 1    | 1.05      | N.D. | 71CD       | Savakis <i>et al.</i> , 1980       |
| EIP 29            | N.D. | 2    | 2.3       | N.D. | 71CD       | Savakis <i>et al.</i> , 1984       |
|                   |      |      | 1.3       | N.D. |            |                                    |
| EIP 40            | N.D. | 3    | 1.8       | N.D. | 55BE       |                                    |
|                   |      |      | 1.65      | N.D. |            |                                    |
|                   |      |      | 1.5       | N.D. |            |                                    |
| hsp 22            | 1.03 | 1    | 1.03      | 0    | 67B        | Ingolia and Craig, 1982            |
| hsp 23            | 0.96 | 1    | 0.96      | 0    | 67B        | Berger, 1984                       |
| hsp 26            | 1.05 | 1    | 1.05      | 0    | 67B        |                                    |
| hsp 27            | 1.25 | 1    | 1.25      | 0    | 67B        |                                    |
| hsp 68            | 2.1  | 1    | 2.1       | 0    | 95D        | Holmgren <i>et al.</i> , 1979      |
| hsp 70            | 2.25 | 1    | 2.25      | 0    | 87A7, 87C1 |                                    |
| hsp 83            | 3.7  | 1    | 2.75      | 1    | 63BC       |                                    |
| hsp 93            | N.D. | N.D. | N.D.      | N.D. | 93D        | Walldorf <i>et al.</i> , 1984      |

|                  |      |            |                           |      |           |   |
|------------------|------|------------|---------------------------|------|-----------|---|
| Histone 1        | 0.95 | 1          | 0.95                      | 0    | 39DE      | Lifton <i>et al.</i> , 1977   |
| Histone H2a      | 0.47 | 1          | 0.47                      | 0    | 39DE      | Pardue <i>et al.</i> , 1977   |
| Histone H2b      | 0.42 | 1          | 0.42                      | 0    | 39DE      | Saigo <i>et al.</i> , 1980  |
| Histone H3       | 0.52 | 1          | 0.52                      | 0    | 39DE      |   |
| Histone H4       | 0.36 | 1          | 0.36                      | 0    | 39DE      |   |
| Tropomyosin-1    | 4.3  | 4          | 1.9                       | 4    | 88F2-5    | Basi <i>et al.</i> , 1984   |
|                  |      |            | 1.7                       | 4    |           | Bautch <i>et al.</i> , 1982   |
|                  |      |            | 1.6                       | 3    |           | Karlik <i>et al.</i> , 1984   |
|                  |      |            | 1.3                       | 3    |           |   |
| Tropomyosin-2    | 2.3  | 1          | 1.5                       | N.D. | 88F2-5    |   |
| Tropomyosin-3    | 3.0  | N.D.       | N.D.                      | N.D. | 88F2-5    | Bautch and Storti, 1983   |
| b) Complex       |      |            |                           |      |           |   |
| Notch            | 40   | 4          | 10.5                      | 8    | 3C7       | Artavanis-Tsakonas <i>et al.</i> , 1983   |
|                  |      |            | 7.0                       | N.D. |           | Kidd <i>et al.</i> , 1983   |
|                  |      |            | 0.9                       | N.D. |           |   |
|                  |      |            | 0.7                       | N.D. |           |   |
| rudimentary      | 13.2 | 1          | 7.3                       | 3    | 15A1      | Segraves <i>et al.</i> , 1983; Segraves <i>et al.</i> , 1984                          |
| scute            | 60.0 | 4          | N.D.                      | N.D. | 1B        | Carramolino <i>et al.</i> , 1982  |
| white            | 6.0  | at least 2 | 2.6                       | 4    | 3C        | O'Hare <i>et al.</i> , 1983; Pirrotta <i>et al.</i> , 1983; Pirrotta and Brockl, 1984 |
|                  |      |            | 2.2                       | N.D. |           | Fjose <i>et al.</i> , 1984  |
| c) Developmental |      |            |                           |      |           |   |
| i) homeotic      |      |            |                           |      |           |   |
| bithorax         | 73   | 3          | 4.7 (polyA <sup>-</sup> ) | 2?   | 89E1-89E4 | Bender <i>et al.</i> , 1983   |
|                  |      |            | 4.3                       | 3    |           | R. Saint, P. Beachy, David Hogness (manuscript in preparation)                        |
|                  |      |            | 3.0                       | 3    |           |   |
| Antennapedia     | 100  | 3          | 5.0                       | 3?   | 84AB      | Scott <i>et al.</i> , 1983  |
|                  |      |            | 3.5                       | N.D. |           | Garber <i>et al.</i> , 1983   |
|                  |      |            | 1.1 (polyA <sup>-</sup> ) | N.D. |           |   |

ii) pattern determination

|                      |        |     |      |      |          |  |
|----------------------|--------|-----|------|------|----------|--|
| engrailed            | 40     | 3   | 3.6  | N.D. | 48A      | Fjose <i>et al.</i> , 1985                                   |
|                      |        |     | 2.7  | 2    |          | Poole <i>et al.</i> , 1985                                   |
|                      |        |     | 1.4  | N.D. |          |  |
| fushi tarazu         | 2.02   | 1   | 1.8  | 1    | 84AB     | Kuroiwa <i>et al.</i> , 1984;<br>Weiner <i>et al.</i> , 1984 |
| Kruppel              | 4.0    | 1   | 2.5  | N.D. | 60F2-5   | Priess <i>et al.</i> , 1985                                  |
| iii) maternal effect |        |     |      |      |          |  |
| dorsal               | 8.0    | 1?  | 2.8  | N.D. | 36C      | Steward <i>et al.</i> , 1984                                 |
| d) Neurobiological   |        |     |      |      |          |  |
| dopa decarboxylase   | 4.0    | 5?  | 4.0  | N.D. | 37C1     | Hirsh and Davidson, 1984                                     |
|                      |        |     | 3.0  | N.D. |          | Beall and Hirsh 1984   |
|                      |        |     | 2.7  | N.D. |          |  |
|                      |        |     | 2.3  | N.D. |          |  |
|                      |        |     | 2.0  | 2    |          |  |
| dunce                | 26     | 6   | N.D. | N.D. | 3C12-3D4 | Davis and Davidson, 1984;<br>R. Davis, <i>pers. comm.</i>    |
| period               | 19-20? | 5.6 | 4.5  | N.D. | 3B       | Reddy <i>et al.</i> , 1984                                   |
|                      |        |     | 3.2  | N.D. |          | Bargiello and Young, 1984                                    |
|                      |        |     | 2.7  | N.D. |          |  |
|                      |        |     | 1.7  | N.D. |          |  |
|                      |        |     | 1.0  | N.D. |          |  |
|                      |        |     | 0.9  | N.D. |          |  |



Table 1.2

Genes potentially located in divisions 19 and 20 of *D. melanogaster*

| <u>Mutant</u>                      | <u>Chromosome Position</u> | <u>Phenotype</u>   | <u>Reference</u>                     |
|------------------------------------|----------------------------|--|--------------------------------------|
| a) DEVELOPMENTAL                   |                            |  |                                      |
| tumorous-head -1 ( <i>tuh-1</i> )  | X:65.3                     | Eyes, rostralhaut and antennae are changed to abdominal, genital and leg structures, respectively.   | Woolf and Passage, 1980              |
| runt ( <i>run</i> )                | X:65.0                     | Some segments absent, anterior portion of remaining segments is duplicated   | Nusslein-Volhard and Wieschaus, 1980 |
| folded gastrulation ( <i>fog</i> ) | 20A3-20B <sup>(a)</sup>    | Gastrulation defective, extensive folds in the germ band, ventral holes in cuticle, particularly at anterior and posterior ends                                | Wieschaus <i>et al.</i> , 1984       |
| extra organs ( <i>eo</i> )         | 20A1-2 <sup>(a)</sup>      | Eyes and wings malformed, antennae or arista duplicated or triplicated, leg branched or duplicated   | Schalet, 1972                        |
| b) VISUAL SYSTEM                   |                            |  |                                      |
| small optic lobes ( <i>sol</i> )   | X:60                       | Medulla, lobula and lobula plate optic lobes reduced in volume and cell number, electroretinogram lacks lamina potentials, poor orientation to spots in Y maze | Bulthoff, 1982                       |
| EE-111                             | X:65.0                     | Weak phototaxis in counter-current apparatus; more photopositive than wildtype in maze test  | Markow and Merriam, 1977             |
| PC-16                              | X:65.0                     | Poor phototaxis in counter-current apparatus; more photopositive than wildtype in maze test; not allelic to EE-111   | Markow and Merriam, 1977             |
| c) FLIGHT AND THORACIC MUSCLES     |                            |  |                                      |
| flightless-H ( <i>flh</i> )        | X:60.0                     | Flightless; no wing beating; jumping ability poor; mosaic analysis suggests primary defect could be in CNS or in muscles                                       | Homyk and Sheppard, 1977             |
| flight-reduced-L ( <i>flrdL</i> )  | X:62.0                     | Weak flight; generally inactive  | Homyk and Sheppard, 1977             |
| flightless-1 ( <i>flil</i> )       | X:65.0                     | Flight impaired; jumping ability poor  | Homyk and Sheppard, 1977             |
| flight-defective 0 ( <i>flt0</i> ) | 19F 20F <sup>(a)</sup>     | Completely flightless, mosaic analysis suggests primary defect is in thoracic muscles  | Deak <i>et al.</i> , 1982            |

d) PHYSIOLOGICAL AND  
MOTOR COORDINATION

|   |                          |  |   |
|---|--------------------------|--|---|
| jumpless ( <i>jpl</i> <sup>4229</sup> ) | X:62.1                   | Unable to jump or fly; dorsolongitudinal flight muscles of thorax not driven by giant nerve fibre or associated inter-neuron or both   | Hall, 1982                                  |
| EE-85                                   | X:65                     | Inexcitable in phototaxis test; sluggish in general locomotion; mosaic analysis maps sluggishness to head-thoracic region  | Markow and Merriam, 1977                    |
| Shaking-B ( <i>ShakB</i> )              | X:65                     | Adults hyperactive; after etherisation appendages move in uncoordinated fashion  | Homyk <i>et al.</i> , 1980                  |
| stress-sensitive-C ( <i>sesC</i> )      | X:65                     | Paralysed briefly by mechanical stress, followed by sporadic movement; jumps and flies abnormally short distances  | Homyk and Shepphard, 1977                   |
| uncoordinated ( <i>unc</i> )            | X:65.5                   | Unable to walk because of lack of coordinated leg movements; wings held up and frequently curled at tips; viability reduced; death occurs soon after eclosion                              | Schalet, 1972                               |
| uncoordinated-like ( <i>uncl</i> )      | X:66.0                   | Uncoordinated leg movements; death occurs soon after eclosion; majority of mutants die just before eclosion  | Schalet, 1972                               |
| stoned ( <i>stn</i> <sup>ts</sup> )     | X:66.3                   | Aberrant locomotion at low temperatures; immobility induced on shift to 29°C; causes abnormally accentuated phosphorylation of certain proteins in synaptosomal fractions from adult heads | Grigliatti <i>et al.</i> , 1973; Hall, 1982 |
| passover ( <i>pas</i> )                 | 19D1-20A2 <sup>(a)</sup> | Lacks direct transmission between the giant fibre and the jump motor neuron due to abnormality of the jump motor neuron  | Wyman and Thomas, 1983                      |
| e) MEMORY                               |                          |  |   |
| amnesiac ( <i>am</i> )                  | X:60.0                   | Learns normally but forgets four times as quickly as wildtype flies; probably affects memory retrieval rather than storage   | Quinn <i>et al.</i> , 1979                  |

(a) Chromosome position determined cytogenetically rather than by recombination analysis.

Figure 1.1: Comparison of the mitotic, polytene and genetic recombination maps of the X chromosome.

a) Schematic diagram of the mitotic X chromosome, showing the major constrictions. The heterochromatic region is shaded and the euchromatin unshaded. The four centromeric regions are designated hA, hB, hC and hD. The kinetochore is subterminal and the genetic right hand limb is designated XR. NO indicates the nucleolus organizer.

b) Salivary gland polytene chromosome with each division marked.

c) Genetic recombination map. (Adapted from Hannah, 1951).

ation

major

matin

The

d XR.

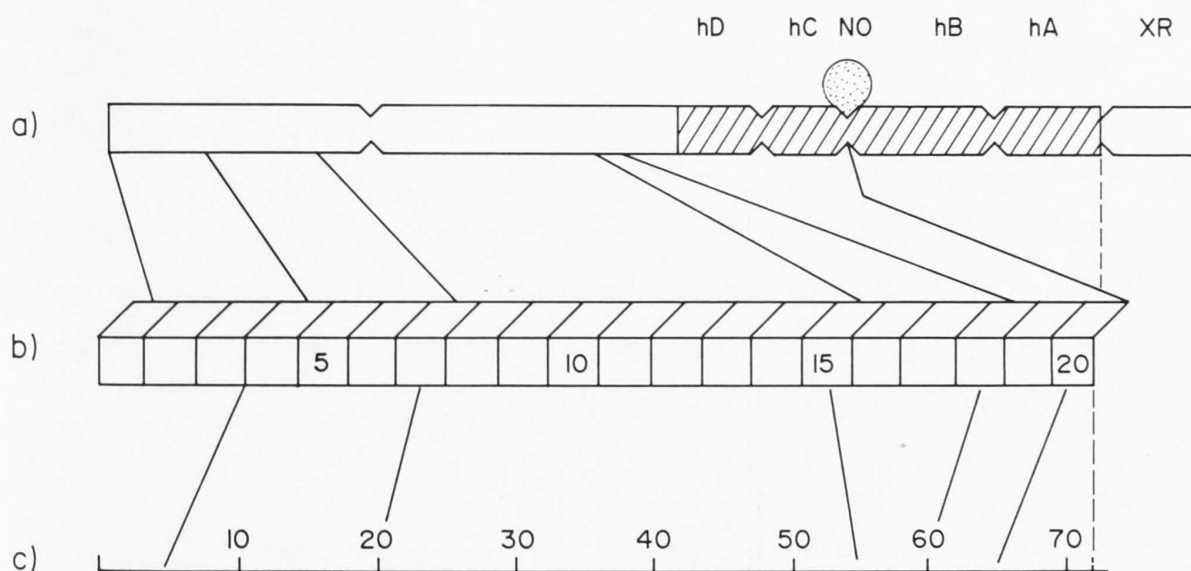


Figure 1.2: The distribution of radiation induced chromosomal rearrangements along the X chromosome.

X-ray induced lethal X chromosomes were randomly recovered and the breakpoints of the rearrangements determined. The subdivisions of the polytene chromosomes studied are indicated. The heterochromatin is illustrated schematically (Summarized from Lefevre, 1981).



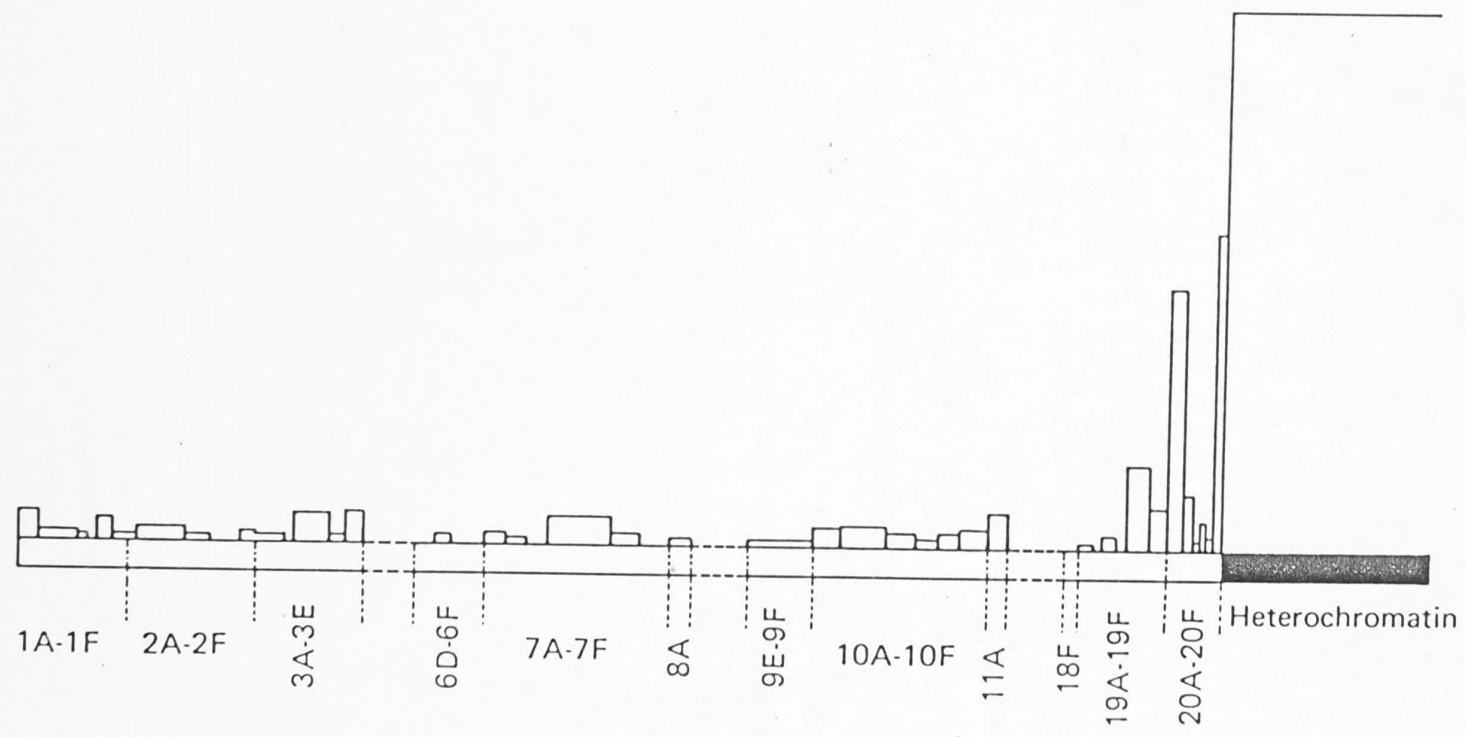
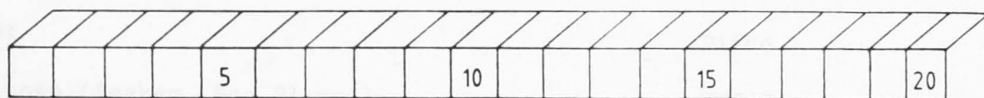


Figure 1.3: The proximal region of the X chromosome.

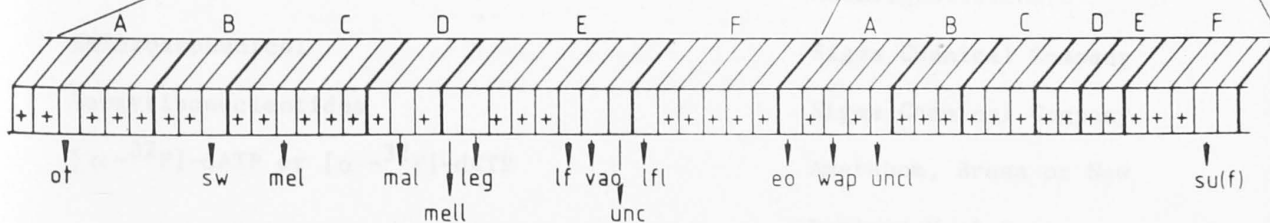
a) Schematic diagram of the polytene X chromosome showing the chromosomal divisions.

b) Divisions 19 and 20 illustrated in detail, showing the subdivisions (lettered) and complementation groups where lethal (+) and visible mutations have been identified. The visible mutations are *outheld* (*ot*), *shortwing* (*sw*), *melanized* (*mel*), *maroon-like* (*mal*), *melanized-like* (*mell*), *legless* (*leg*), *little-fly* (*lf*), *varied outspread* (*vao*), *uncoordinated* (*unc*), *little-fly-like* (*lfl*), *extra organs* (*eo*), *wings apart* (*wap*), *uncoordinated-like* (*uncl*) and *suppressor-of-forked* (*su(f)*). (Adapted from Schalet and Lefevre, 1976).

a)



b)



## CHAPTER TWO

MATERIALS AND METHODS2.1 Materials2.1.1 Chemicals and reagents

The source of the more important reagents and chemicals used in the described experiments are listed.

|  |   |
|--|---|
| Agar   | Difco                                     |
| Agarose (Seakem, Sea Plaque)   | FMC Corporation                           |
| Ampicillin   | Sigma Chemical Company                    |
| Acrylamide   | Bio Rad                                   |
| Bactotryptone  | Difco                                     |
| Bovine Serum Albumin (Fraction V)  | Sigma Chemical Company                    |
| Casamino Acids   | Difco                                     |
| Cesium chloride  | Metallgesellschaft                        |
| Chloroamphenicol   | Sigma Chemical Company                    |
| Deoxyribonucleotides   | Sigma Chemical Company                    |
| [ $\alpha$ - <sup>32</sup> P]-dATP or [ $\alpha$ - <sup>32</sup> P]-dCTP | Amersham, Bresa or New<br>England Nuclear |
| Diethylpyrocarbonate (DEPC)  | Sigma Chemical Company                    |
| Dimethylsulphoxide (DMSO)  | Fluka                                     |
| Ficoll   | Pharmacia P-L Biochemicals                |
| Formamide  | Ajax Chemical Company                     |
| Glyoxal  | Fluka                                     |
| Guanidine Thiocyanate  | Fluka                                     |
| N'laurylsarcosine  | Sigma Chemical Company                    |
| $\beta$ -mercaptoethanol   | Sigma Chemical Company                    |

|  |                        |
|--|------------------------|
| N,N'-methylene-bis-acrylamide (bis)          | BioRad                 |
| Mixed Bed Basin (Analytical Grade)           | BioRad                 |
| Nitrocellulose BA85 (0.45um)                 | Schleicher & Schuell   |
| NZamine (casein enzymatic hydrolysate)       | Sigma Chemical Company |
| Oligo (dT)-cellulose                         | Sigma Chemical Company |
| Phenol (buffer saturated before use)         | Wako Pure Chemicals    |
| Polyvinylpyrrolidone                         | Sigma Chemical Company |
| Ribonucleotides                              | Sigma Chemical Company |
| $\alpha$ - <sup>32</sup> P-GTP               | Bresa                  |
| Sephadex G-50                                | Pharmacia              |
| Spectinomycin-HCl (Trobicin)                 | Upjohn                 |
| N,N,N',N'-tetramethylethylenediamine (TEMED) | BioRad                 |
| Tetracycline                                 | Sigma Chemical Company |
| X-ray film (XPR-5 and XRP-1)                 | Kodak                  |
| Yeast extract                                | Difco                  |
| Zetaprobe                                    | BioRad                 |

All other chemicals and reagents used were analytical grade.

### 2.1.2 Enzymes

The restriction endonucleases *Bam*HI, *Eco*RI, *Hind*III, *Hae*III and *Pst*I were a gift from J. Blok and A. McKenzie. All other restriction enzymes were obtained from New England Biolabs Inc., New England Nuclear or Boehringer Mannheim. DNA Polymerase I large fragment (Klenow) was obtained from New England Biolabs, New England Nuclear or Bresa and SP6 polymerase from Bresa. T4 DNA ligase was a gift from A. McKenzie. Ribonuclease-free deoxyribonuclease I (DNA'ase I) was purchased from Worthington Biochemical Co. and calf intestinal alkaline phosphatase was obtained from Collaborative Research and lysozyme from Sigma Chemical Company. Human placental ribonuclease inhibitor was obtained from Amersham.



### 2.1.3 Bacterial strains and media

*E. coli* strains were stored as suspensions in 50% glycerol at  $-20^{\circ}\text{C}$ . The relevant characteristics of the two host *E. coli* strains used are listed in Table 2.1.

The compositions of the media used for bacterial culture were as follows:

#### LM Broth (Miller, 1972)

|   |            |
|---|------------|
| Bactotryptone                                       | 10g        |
| Yeast extract                                       | 5g         |
| NaCl  | 10g        |
| Maltose   | 2g         |
| $\text{MgCl}_2 \cdot 2\text{H}_2\text{O}$           | 0.2g       |
| Tris-HCl, pH 7.0                                    | 10mM       |
| Double distilled water (D.D. $\text{H}_2\text{O}$ ) | to 1 litre |

#### NZY Broth (Blattner *et al.*, 1977)

|   |            |
|---|------------|
| NZamine                                   | 10g        |
| Yeast extract                             | 5g         |
| NaCl                                      | 5g         |
| $\text{MgCl}_2 \cdot 2\text{H}_2\text{O}$ | 2g         |
| D.D. $\text{H}_2\text{O}$                 | to 1 litre |

NZCY:- as above with the addition of casamino acids (1g/l).

The concentration of antibiotics in media selective for the growth of *E. coli* strains carrying plasmids were as follows, unless otherwise stated:

|                       |         |
|-----------------------|---------|
| Ampicillin (Ap)       | 50ug/ml |
| Tetracycline (Tc)     | 20ug/ml |
| Chloroamphenicol (Cm) | 30ug/ml |

Solid media contained 1.1% agar in all cases.

#### 2.1.4 DNA vectors

Cloning vehicles used in this work are listed in Table 2.2, together with their relevant characteristics and source.

#### 2.1.5 Ribonuclease free glassware and solutions

All glassware required for experiments involving RNA were rendered ribonuclease free by baking for 16 to 18 hours at 200°C. Disposable microfuge tubes and pipetteman tips were autoclaved. Where possible, solutions were treated with 0.1% DEPC overnight at room temperature followed by autoclaving (120°C, 20p.s.i.) (Maniatis *et al.*, 1982). In cases where autoclaving was not possible, solutions were prepared in ribonuclease free glassware using double distilled water that had been DEPC-treated and autoclaved.

#### 2.1.6 Drosophila stocks and media

Table 2.3 lists the wildtype strains and mutant and rearranged X chromosomes used in this study, together with their relevant characteristics and origin. The media used for culturing *Drosophila* was composed of:

|                             |            |
|-----------------------------|------------|
| agar                        | 10g        |
| sucrose                     | 15g        |
| yeast                       | 35g        |
| malt                        | 40g        |
| high protein powder (HyPro) | 10g        |
| corn syrup                  | 30mls      |
| propionic acid              | 4.5ml      |
| ethanol                     | 4.5mls     |
| D.D. H <sub>2</sub> O       | to 1 litre |

## 2.2 Nucleic Acid Preparative Procedures

### 2.2.1 Genomic DNA

Genomic DNA was extracted from adult heads or whole pupae by the procedure of Miklos (1984). Briefly, the adult heads and body parts of organisms that had been snap frozen in liquid nitrogen and stored at  $-90^{\circ}\text{C}$  were fractionated over mesh sieves. The adult heads, or whole pupae, were homogenized in buffer (10mM Tris, pH 8.0, 20mM EDTA) and lysed with the detergent Sarkosyl NL-30 at a final concentration of 3%. Cesium chloride and ethidium bromide were added to the lysate at final concentrations of 1g/ml and 0.6mg/ml, respectively. The DNA was brought to equilibrium in the cesium chloride gradient by ultracentrifugation at 44,000 rpm (Type 50Ti rotor) for 40 to 44 hours, and visualized with U.V. light (366nm). Genomic DNA was withdrawn by side puncture, using a 19 guage needle, and the ethidium bromide was extracted at least six times with isopropanol. The DNA was dialysed extensively against T.E. buffer (10mM Tris-HCl, pH 7.4, 1mM EDTA) at  $4^{\circ}\text{C}$  and precipitated by 0.3M sodium acetate (NaAc) and 2.5 volumes of ethanol at  $-20^{\circ}\text{C}$  overnight. Pelleted DNA was redissolved in T.E. buffer at a concentration of 25 to 50 ug/ml, as determined by comparison to a DNA of known concentration, and visualized with U.V. light in the presence of 1.5ug/ml ethidium bromide.

### 2.2.2 Supercoiled plasmid DNA

Selective broth media were inoculated with an overnight culture of *E. coli* carrying recombinant plasmids (1:100 dilution) and incubated at  $37^{\circ}\text{C}$  with shaking until the optical density at 640nm (O.D.<sub>640</sub>) reached 0.4 to 0.6. Production of the plasmid DNA was amplified by the addition of solid spectinomycin (100 ug/ml), followed by incubation for a further 16 to 18 hours.

Bacterial cells were harvested by centrifugation at 5,000 rpm for 20 minutes at 4°C (GSA rotor), resuspended in 20% sucrose in T.E. buffer (0.8ml/100ml culture) and maintained on ice. Lysozyme, freshly prepared in T.E. buffer, was added to a final concentration of 2.5mg/ml and the cell suspension incubated for 5 minutes. After the addition of 0.7ml of 0.25M EDTA, and a further 2 minute incubation, the cells were lysed with the addition of an equal volume of lytic mix (10% Triton X-100, 1.0M Tris, pH 8.0, 0.25M EDTA). After 10 minutes at 4°C, cell debris was pelleted by centrifugation at 16,500 rpm (SS34 rotor) for 40 minutes, and cesium chloride and ethidium bromide were then added to the cleared lysate at concentrations of 1g/ml and 0.6mg/ml, respectively. Plasmid DNA was brought to equilibrium in the cesium chloride gradient by ultracentrifugation, recovered and dialysed as described in Section 2.2.1. The DNA concentration was determined spectrophotometrically by optical density readings at 260nm, and the purity assessed by a comparison of readings at 260nm and 280nm (Maniatis *et al.*, 1982). If the DNA concentration was less than 50ug/ml, it was precipitated as described in section 2.2.1 and redissolved at a concentration of 50 to 100ug/ml.

### 2.2.3 Bacteriophage DNA

Bacteriophage suspensions at a titre of  $10^3$  to  $10^4$  p.f.u./200ul SM (0.05M Tris-HCl, pH 7.4, 0.1M NaCl, 0.01M MgCl, 0.01% gelatin) were mixed with 200ul of stationary phase *E. coli* LE392 cells and 200ul Mg-Ca solution (0.01M MgCl<sub>2</sub>, 0.01M CaCl<sub>2</sub>). Adsorption of bacteriophage particles was allowed to proceed at 37°C for 10 minutes, after which the cell suspension was inoculated into 100ml NZY or NZCY broth and incubated at 37°C for 16 to 18 hours, or until lysis had occurred. Bacterial cells were recovered by centrifugation at 5,000 rpm for 20 minutes (GSA rotor) and the phage particles precipitated by further

centrifugation at 13,000 rpm for 3 hours (SS34 rotor). Bacteriophage pellets were resuspended in 4.0ml of cold T.E. buffer containing 3.4g cesium chloride and centrifuged at 35,000 rpm for 16 to 18 hours at 20°C (SW55Ti or SW50.1 rotor). The phage band was harvested by side puncture using a 19 gauge needle.

DNA was extracted from the bacteriophage particles by two different methods. One procedure was that of Davis *et al.* (1980). Briefly, 0.1 volumes 2M Tris-HCl, pH 8.5 and an equal volume of deionized formamide were mixed with the bacteriophage suspension in CsCl. After two hours incubation at room temperature, 1 volume of distilled water and 15ug carrier *E. coli* tRNA were added. The DNA was precipitated by the addition of 6 volumes ethanol (room temperature) and centrifugation for 5 minutes in a bench microfuge. Each pellet was rinsed in 70% ethanol, dried *in vacuo* and resuspended in T.E. buffer at a concentration of 50 to 100ug/ml.

Bacteriophage DNA was also extracted according to the procedure of Blattner *et al.* (1977). The bacteriophage suspension in CsCl was dialysed extensively against T.E. buffer at 4°C. An equal volume of 0.1M EDTA was added and the suspension heated at 60°C for 10 minutes. Pronase, that had been previously autodigested at 37°C for 1 hour, and sodium dodecyl sulphate (SDS) were added to final concentrations of 1mg/ml and 0.05%, respectively, and the suspension incubated at 37°C for one hour. The DNA was twice extracted with an equal volume of phenol and chloroform, dialysed extensively against T.E. buffer and precipitated by 0.3M NaAc and 2.5 volumes of ethanol at -20°C (overnight) or -90°C (two hours). The DNA was recovered as described in section 2.2.2.



#### 2.2.4 Total cellular RNA

RNA was isolated from Canton-S mid embryos, early larvae, late larvae, mid pupae and adults, cultured and harvested as described in section 2.8.2. RNA was also extracted from homozygous *uncoordinated* mutants and organisms carrying deficiency chromosomes, which were collected as late imagos and newly eclosed adults, and adult heads, which were obtained as described in section 2.2.1. All organisms selected for RNA extractions were snap frozen in liquid nitrogen and stored at  $-90^{\circ}\text{C}$ .

Total cellular RNA was prepared by a modification (A.J. Howells, *pers. comm.*) of the method of Chirgwin *et al.* (1979). Whole organisms were finely ground in a mortar and pestle which had been pre-cooled with liquid nitrogen. Approximately 8mls extraction solution (4M guanidine thiocyanate, 50mM sodium acetate, pH 5.5, 1M  $\beta$ -mercaptoethanol, extracted with acid washed activated charcoal and filtered, plus 5mM EDTA, pH 7.5, and 1% N-lauryl sarcosine) per gram of tissue and solid cesium chloride (0.5g/ml extraction solution) were added, and the mortar and pestle heated at  $60^{\circ}\text{C}$ . When all solid material had dissolved, the preparation was quick-cooled on ice. Four ml of the RNA solution was layered onto a 1.2ml cesium chloride cushion solution (5.7M cesium chloride, 0.01M Tris-HCl, pH 7.5, 0.1M EDTA, pH 7.5) and centrifuged at 35,000 rpm for 16 to 18 hours at  $15^{\circ}\text{C}$  (SW 50.1 or SW 55Ti rotor).

The RNA pellet was resuspended in T.E. buffer, pH 7.5, plus 0.5% N-lauryl sarcosine, heated at  $80^{\circ}\text{C}$  for 1 minute and quick-cooled on ice. The solution was extracted with an equal volume of phenol/chloroform/isoamyl alcohol (25:25:1) and the RNA precipitated by 0.3M NaAc and three volumes of ethanol at  $-20^{\circ}\text{C}$  overnight. The RNA was recovered by centrifugation at 10,000 rpm for 30 minutes (SS34 rotor). The pellet was rinsed with 70% ethanol, dried *in vacuo* and redissolved in double distilled water. The concentration and purity

of the RNA preparations were determined as described in section 2.2.2, after which the RNA was dispensed into small aliquots and stored under ethanol at  $-20^{\circ}\text{C}$ .

#### 2.2.5 Polyadenylated RNA

Polyadenylated RNA was fractionated from the total cellular RNA population by affinity column chromatography as described by Maniatis *et al.* (1982). The chromatography column was prepared by equilibrating lg oligo (dT)-cellulose in loading buffer (20mM Tris-HCl, pH 7.6, 0.5M NaCl, 1mM EDTA, 0.1% SDS) and packing a 3ml column under gravity. The column was washed with 3 volumes each of distilled water, 0.1M NaOH, 5mM EDTA and distilled water again, followed by 5 volumes of loading buffer.

The RNA samples, stored under ethanol, were recovered by centrifugation for 10 minutes in a bench microfuge and dried *in vacuo*. The RNA was redissolved in distilled water, heated at  $65^{\circ}\text{C}$  for 5 minutes and an equal volume of twice-concentrated (2X) loading buffer was added. The sample was cooled to room temperature and applied to the column under gravity. The flow-through was collected, again heated at  $65^{\circ}\text{C}$  for 5 minutes, cooled and reapplied to the column. The column was washed with 10 volumes of loading buffer (with the flow rate adjusted to 1ml per minute using an LKB peristaltic pump), followed by 4 volumes of loading buffer containing 0.1M NaCl. Fractions of approximately 5ml were collected and the optical density at 260nm determined to ensure that the last fractions contained very little material that absorbed at this wavelength.

The polyadenylated RNA was eluted from the column with 3 volumes of elution buffer (10mM Tris-HCl, pH 7.5, 1mM EDTA and 0.05% SDS). The poly A<sup>+</sup> RNA was precipitated by 0.3M NaAc and 3 volumes of ethanol at  $-90^{\circ}\text{C}$  overnight. The RNA was recovered as described in section 2.2.4 and the

concentration and purity determined as described in section 2.2.2. The poly A<sup>+</sup> RNA was dispensed into aliquots of 6 $\mu$ g and stored under ethanol at -20°C. Typically, 2mg of total cellular RNA was applied to the column and a yield of 2 to 5% poly A<sup>+</sup> RNA was obtained.

## 2.3 Preparation of Nucleic Acids for Electrophoresis

### 2.3.1 Preparation of DNA samples

DNA samples were analysed by electrophoresis after cleavage by restriction endonucleases. Single restriction enzyme digestion reactions were performed in accordance with the manufacturer's directions. However, digestions involving *Hind*III, and those employing more than one enzyme, were usually carried out in T.A. buffer (33mM Tris-HCl, pH 7.9, 66mM potassium acetate, 10mM magnesium acetate, 5mM dithiothreitol (DTT) (O'Farrell *et al.*, 1980)). For analytical purposes, 0.5 to 1.0 $\mu$ g DNA was cleaved by incubation with 2 or more units of enzyme in a final volume of 22 to 25 $\mu$ l at 37°C for 1 to 2 hours. The reaction was stopped by the addition of 5 $\mu$ l sample dye (30% sucrose, 0.09% bromophenol blue, 50mM EDTA). Preparative digestion reactions (>5 $\mu$ g DNA) were performed in a final volume of 100 to 200 $\mu$ l for 2 hours with a proportional increase in the amount of enzyme used. The reaction was stopped by either heat treatment at 65°C for 5 minutes or by the addition of EDTA (to 10mM) followed by extraction with an equal volume of phenol and chloroform. DNA was precipitated with ethanol as described in section 2.2.1.

### 2.3.2 Preparation of RNA samples

RNA samples were prepared for electrophoresis by denaturation with glyoxal and DMSO, as described by McMaster and Carmichael (1977) and Carmichael and McMaster (1980). Unless otherwise indicated, 30 $\mu$ g of total cellular RNA and 6 $\mu$ g of polyadenylated RNA were tested per sample.

RNA samples, stored under ethanol, were recovered after centrifugation for 10 minutes in a bench microfuge and dried *in vacuo*. Each pellet was dissolved in 5ul T.E. buffer, pH 7.5, and denatured in the presence of 50% deionized DMSO, 6% deionized glyoxal and 10mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{PO}_4$  buffer, pH 7.0, in a total volume of 20ul. DMSO and glyoxal were deionized just prior to the reaction by at least five passages through a small column of mixed bed resin. Following incubation at 50°C for 1 hour, the samples were cooled on ice and the reaction stopped by the addition of 2ul sample dye (50% glycerol, 2.5 mM phosphate buffer, pH 7.0, 0.09% bromophenol blue).

## 2.4 Electrophoresis of Nucleic Acids

### 2.4.1 Agarose gel electrophoresis of DNA

Electrophoresis of restricted DNA fragments in the 23 to 0.7kbp size range was performed in horizontal 1% agarose gels (16cm x 17cm x 1cm) prepared in Tris-acetate buffer (0.04M Tris-HCl, pH 7.8, 0.005M sodium acetate, 0.001M EDTA) containing 1.5ug/ml ethidium bromide. Electrophoresis was generally carried out at 1.5 to 2.5V/cm for 16 to 18 hours at room temperature. Lambda cl857 DNA, cleaved with *Hind*III, provided size markers. DNA bands were visualized by transmitted short wave (254nm) U.V. light and recorded on Polaroid Type 55 Land Positive/Negative Film.

Rapid analysis of restricted DNA and the preparative separation of DNA fragments were performed in small horizontal agarose gels (9cm x 5.5cm x 0.5cm) prepared in Tris-borate buffer (0.09M Tris, 0.08M boric acid, 2.5M EDTA, pH 8.3) for 2 to 4 hours at 8V/cm. Preparative electrophoresis was performed in low melting point (Sea Plaque) agarose. Gels were stained with 10ug/ml ethidium bromide for 15 minutes and the DNA visualized by transmitted short wave (254nm) U.V. light for analytical purposes, or reflected long wavelength (366nm) U.V. light in the case of preparative samples.

DNA fragments required for radio-labelling or cloning were electroeluted from agarose or acrylamide gels. Excised gel slices were placed in dialysis tubing containing 0.5X Tris-borate buffer, and the DNA electroeluted at 80V, (20mA) for 2 hours. The eluate was extracted with an equal volume of phenol and chloroform and the DNA precipitated by 0.3M NaAc and 2.5 volumes of ethanol at  $-90^{\circ}\text{C}$  for at least 2 hours.

#### 2.4.2 Polyacrylamide gel electrophoresis of DNA

Vertical slab gels (10% polyacrylamide) were used for the resolution of DNA fragments of 50pb to 700bp. Gels (140mm x 160mm x 1.5mm) were prepared by 50% dilution of a 20% acrylamide, 1% bis stock solution (filtered, degassed and stored at  $4^{\circ}\text{C}$ ) with Tris-borate buffer. Polymerization of the acrylamide was achieved using 0.1% ammonium persulphate and 0.6% TEMED. Electrophoresis was carried out at 14V/cm for 4 to 6 hours in Tris-borate buffer. Gels were stained and the DNA visualized and recorded as described in section 2.4.1. Molecular size markers were provided by *Hinf*I digestion of plasmid pBR322 DNA.

#### 2.4.3 Agarose gel electrophoresis of RNA

RNA fragments were size fractionated in horizontal 1% agarose gels (19cm x 14cm x 0.7cm) prepared in sodium phosphate buffer (10mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{PO}_4$ , pH 7.0) (McMaster and Carmichael, 1977; Carmichael and McMaster, 1980). Electrophoresis was carried out at 1.5 V/cm for 16 hours at room temperature, with buffer recirculation (1200mls per hour). Molecular size markers were provided by *Hind*III digestion of lambda c1857 DNA. After cleavage, the size markers were extracted with an equal volume of phenol and chloroform and precipitated by 0.3M NaAc and 2.5 volumes of ethanol at  $-90^{\circ}\text{C}$  for at least 2 hours. Markers were denatured for electrophoresis as described in section 2.3.2.



## 2.5 Molecular Cloning Protocol

### 2.5.1 Preparation of DNA samples

Restriction enzyme cleavage of DNA in preparation for cloning was as described in section 2.3.1.

### 2.5.2 Dephosphorylation of vectors

The efficiency of vector self-ligation was drastically reduced by dephosphorylation of the terminal 5' phosphates of digested vector DNA using calf-intestinal alkaline phosphatase (CAP). DNA (10ug) was treated with CAP (2 units) in the presence of 100mM Tris base, pH 10.8 and 0.2% SDS, for 1 hour at 37°C. The reaction was stopped with two successive extractions with an equal volume of phenol and chloroform, and the DNA precipitated with 0.3M NaAc and 2.5 volumes of ethanol at -20°C overnight. The vector DNA was recovered by centrifugation for 15 minutes in a bench microfuge, washed with 70% ethanol, dried *in vacuo* and redissolved in T.E. buffer at a concentration of 0.5ug/1ul. The efficiency of dephosphorylation was assessed by transforming competent RRI bacterial cells (see Section 2.5.4) with digested, dephosphorylated vector DNA and noting the number of transformed cells obtained under selective growth conditions.

### 2.5.3 Ligation conditions

Plasmid or bacteriophage DNA (2ug) previously digested with an appropriate restriction endonuclease, was ligated to 0.5ug dephosphorylated vector DNA in the presence of *Hae*III buffer (6mM NaCl, 6mM Tris-HCl, pH 7.4, 6.6mM MgCl<sub>2</sub>, 6mM  $\beta$ -mercaptoethanol), 1mM ATP (0.1mM ATP for blunt-ended ligations) and T4 DNA ligase (2 units) in a final volume of 20ul. A control reaction without the enzyme was also performed. Ligations were allowed to proceed at 4°C overnight and the extent of the ligation was assessed by subjecting an aliquot of each reaction to agarose gel electrophoresis.



#### 2.5.4 Transformation procedures

Competent RRI cells were prepared as described by Davis *et al.* (1980) and stored in 500ul aliquots at  $-90^{\circ}\text{C}$ . Briefly, log phase cells were harvested by centrifugation at 5,000 rpm for 5 minutes at  $4^{\circ}\text{C}$  (SS34 rotor) and resuspended in 0.25 volumes of cold 0.1M  $\text{MgCl}_2$ . The cells were sedimented by centrifugation as above, resuspended in 0.25 volumes of 0.1M  $\text{CaCl}_2$ , incubated at  $4^{\circ}\text{C}$  for 30 minutes, and again sedimented by centrifugation. Following resuspension in 0.05 volumes of 0.1M  $\text{CaCl}_2$ , 15% glycerol, the cells were chilled at  $4^{\circ}\text{C}$  for 5 minutes, aliquoted into cold microfuge tubes and snap frozen in liquid nitrogen.

Just prior to transformation, an aliquot of competent cells was thawed on ice. Ligated DNA (made up to a volume of 100ul with T.E. buffer) was added to 200ul competent bacterial cells, incubated at  $4^{\circ}\text{C}$  for 30 minutes and heat-shocked at  $42^{\circ}\text{C}$  for 2 minutes. The transformation mix was then inoculated into 4.7ml LM broth and incubated at  $37^{\circ}\text{C}$  for 1 hour. Aliquots of 100ul were plated onto LM media containing appropriate antibiotics at concentrations described in section 2.1.3. Typically, 1 to  $2 \times 10^3$  recombinants/ml were obtained.

#### 2.5.5 Selection of recombinants

Recombinants were selected by assaying for insertional inactivation of antibiotic resistance genes. Bacteria that failed to grow on the test antibiotic were selected and the size of the inserted fragment determined by a modification of the method of Holmes and Quigley (1981). Briefly, 1ml NZY broth cultures were inoculated from a single colony and incubated overnight at  $37^{\circ}\text{C}$ . The bacterial cells were harvested by centrifugation for 10 minutes in a bench microfuge and lysed by the addition of 50ul STET buffer (8% sucrose, 0.05M EDTA, 0.05M Tris-HCl, pH 8.0, 5% Triton X-100) and 2ul freshly prepared

lysozyme (10mg/ml). After boiling for 40 seconds, the cell debris was removed by centrifugation (10 minutes). Plasmid DNA was precipitated from the lysate by the addition of an equal volume of isopropanol and incubation at  $-20^{\circ}\text{C}$  for 30 minutes. DNA was recovered by centrifugation for 5 minutes, washed with 70% ethanol and dried *in vacuo*. Each pellet was dissolved in 50ul double distilled water and 10ul aliquots digested with the appropriate restriction endonucleases as described in section 2.3.1. Digests were analysed by electrophoresis according to section 2.4.1.

## 2.6 Screening Genomic Libraries

Genomic libraries were screened using the plaque filter hybridization technique of Benton and Davis (1977). 1 to  $3 \times 10^4$  recombinant bacteriophage (representing 2 to 4 genome equivalents of DNA) were adsorbed (for 10 minutes at  $37^{\circ}\text{C}$ ) to 1 to 2ml stationary phase *E. coli* LE392 cells in the presence of an equal volume of Mg-Ca solution. Five ml top agarose (0.6% agarose in NZCY broth) were mixed with 200ul aliquots of adsorbed phage, plated onto 9cm NZCY agar plates and incubated overnight at  $37^{\circ}\text{C}$ .

Before transfer of phage to filters, the plates were cooled at  $4^{\circ}\text{C}$  for 1 hour to prevent adherence of the top agarose to the nitrocellulose. Bacteriophage plaques were adsorbed to circular 0.45um nitrocellulose filters and the DNA subsequently denatured *in situ* by soaking in 0.4M NaOH, 0.8M NaCl for 5 minutes. The filters were neutralized in 0.5M Tris-HCl, pH 7.4, 1.5M NaCl for 5 minutes and washed in 2X SSC (1X SSC: 0.15M NaCl, 0.015M tri-sodium citrate) for at least 10 minutes. Filters were then blotted dry and the DNA bound to the nitrocellulose by baking under vacuum at  $80^{\circ}\text{C}$  for 2 to 4 hours. Up to 4 replicate filters were obtained sequentially from each plate. Filters were hybridized, washed and autoradiographed as described in section 2.7.2. Autoradiographs were aligned with the filters, hybridizing plaques identified and transferred to 1.0ml SM using a sterile pasteur pipette.

Bacteriophage were plaque-purified by further platings, followed by transfer and hybridization until all plaques present on a single plate hybridized to the probe. A single plaque was then selected and transferred to 1.0ml SM. High titre phage stocks were prepared by plating sufficient phage to achieve confluence. The top agarose was then scraped into 5.0ml SM and, after an overnight incubation at 4°C, removed by centrifugation (10 minutes at 2,000 rpm in a bench centrifuge). The supernatant represented the high titre stock which was stored at 4°C in the presence of a few drops of chloroform. Titres were determined by spotting 10ul of a serial dilution series of each stock onto a NZCY plate which contained 200ul fresh overnight culture of *E. coli* LE392 in 5ml top agarose. Plaques were then counted after an overnight incubation at 37°C.

## 2.7 Hybridization Procedures

### 2.7.1 Preparation of hybridization probes

Double stranded DNA hybridization probes were prepared by primed synthesis with DNA polymerase I (large fragment), using denatured calf thymus DNA primers (Whitfield *et al.*, 1982). DNA (0.5 to 1.0ug) was digested with the restriction enzyme *Hae*III as described in section 2.3.1. Digested DNA was heat-denatured by boiling for 5 minutes in the presence of 100ng random primers (8-12 nucleotide fraction of DNA'ase I treated calf thymus DNA) and quick cooled on ice. A mixture of deoxyribonucleotides (dATP, dGTP and dTTP at a final concentration of 1mM each), 30 or 40 uCi [  $\alpha$ -<sup>32</sup>P]-dCTP (3,000 Ci/mmole) and 2-5 units DNA polymerase I (large fragment) were added to the DNA in *Hae*III buffer (6mM NaCl, 6mM Tris-HCl, pH 7.4, 6.6mM MgCl<sub>2</sub>, 6mM  $\beta$ -mercaptoethanol) in a reaction volume of 22ul. If [  $\alpha$ -<sup>32</sup>P]-dATP were used, 1mM dCTP was substituted for dATP in the reaction mix. The reaction was allowed to proceed at 37°C for 30 minutes. Two alternative procedures were

used to separate the unincorporated nucleotides from the radiolabelled DNA. The reaction volume was adjusted to 100ul with T.E. buffer and the reaction mix layered onto a disposable 1ml column of Sephadex-G50 which had been equilibrated in T.E. buffer. Radiolabelled DNA was eluted by centrifugation of the column for 10 minutes at 2,000 rpm (bench centrifuge). Alternatively, the reaction volume was increased to 200ul with T.E. buffer and the reaction stopped by extraction with an equal volume of phenol and chloroform. The unincorporated nucleotides were removed by two successive precipitations, in the presence of 25ug carrier *E. coli* RNA, with 2M ammonium acetate and 2 volumes of ethanol at  $-90^{\circ}\text{C}$  for at least 2 hours. The DNA was recovered by centrifugation for 30 minutes at  $4^{\circ}\text{C}$ , washed in 70% ethanol, dried *in vacuo* and redissolved in 400ul sonicated, denatured herring sperm DNA (6mg/ml). Typically, specific activities of  $2 \times 10^7$  to  $10^8$  cpm/ug were obtained.

Single stranded RNA hybridization probes were synthesized using the SP6 transcription system (Krieg and Melton, 1984(a); 1984(b)). Briefly, fragments from which probes were to be made were subcloned into the pSP64/65 vectors, which contain an M-13 polylinker attached directly to the SP6 transcription initiation site. Plasmid DNA (2ug) was linearized by digestion with a restriction endonuclease (as described in section 2.3.1) that cleaved downstream of the inserted fragment. The digestion reaction was terminated by extraction with an equal volume of phenol and chloroform, and the DNA precipitated as described in section 2.2.1. The DNA was recovered by centrifugation at  $4^{\circ}\text{C}$  for 30 minutes (bench microfuge), washed with 70% ethanol and dried *in vacuo*. Radiolabelled RNA was synthesized by incubation ( $37^{\circ}\text{C}$  for 1hour) of the DNA in transcription buffer (40mM Tris-HCl, pH 7.5, 6mM  $\text{MgCl}_2$ , 2mM spermidine), 10mM DTT, 35 units human placental ribonuclease inhibitor, 0.5mM each of ATP, CTP and UTP and 0.01mM GTP, 50 uCi [ $\alpha$ - $^{32}\text{P}$ ]-GTP (1500 Ci/mmol) and 10 units SP6 RNA polymerase in a reaction volume of

20ul. The DNA template was removed at the end of the reaction by the addition of 80ul 10mM TrisHCl, pH7.5, 10mM MgCl<sub>2</sub>, 35 units human placental ribonuclease inhibitor and 2.5ug DNA'ase (RNA'ase-free) for 10 minutes at 37°C in the presence of 3mM DTT. The reaction was stopped by the addition of 10ug carrier yeast tRNA followed by extraction with an equal volume of phenol and chloroform. The unincorporated nucleotides were separated from the radiolabelled RNA by two successive precipitations with ammonium acetate, as described above.

The specific activities of RNA hybridization probes were typically 1 to 4 x 10<sup>7</sup> cpm/ug template DNA.

### 2.7.2 Southern blot hybridizations

DNA was transferred from agarose gels to nitrocellulose by the method of Southern (1975). The DNA was denatured by subjecting gels to two 45 minute washes with 500ml 0.8M NaCl, 0.4M NaOH, and subsequently neutralized by two 60 minute washes with 500ml 1.5M NaCl, 0.5M Tris-HCl, pH 7.4. Denatured DNA was transferred to nitrocellulose by blotting the gel overnight in 20X SSC. The filter was rinsed in 2X SSC and the DNA fixed by baking for 2 to 4 hours at 80°C *in vacuo*.

Several different hybridization protocols were used. In the first (Miklos *et al.*, 1984), filters were prehybridized in 3X SSC containing 10X Denhardt's solution (1X Denhardt's solution: 0.02% ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin; Denhardt, (1966)) and 150 to 200 ug/ml sonicated denatured herring sperm DNA 65°C overnight. Hybridizations were performed in the prehybridization solution diluted 1:10 with 3X SSC. Radioactively labelled DNA, recovered by passage through Sephadex G50 (section 2.7.1) was denatured by boiling and added to the hybridization solution containing the nitrocellulose filters.



Hybridizations were allowed to proceed at 65°C for 4 hours. Filters were then washed extensively with 2X SSC at 65°C, dried and exposed to X-ray film at -90°C using DuPont Lightning Plus intensifying screens.

An alternative hybridization protocol was used to analyse genomic DNA isolated from *uncoordinated* mutant alleles (section 5.1) (K. Reed *pers. comm.*). Nitrocellulose filters were pre-hybridized in 5X SSPE buffer (1X SSPE:0.18M NaCl, 0.01M phosphate buffer, pH 7.0, 1mM EDTA), 5X Denhardt's solution, 0.1% SDS and 500ug/ml sonicated, denatured herring sperm DNA at 65°C for 4 to 8 hours. Hybridizations were carried out in a freshly prepared prehybridization solution containing only 250 ug/ml carrier DNA. Radioactive probes recovered by ammonium acetate/ethanol precipitation (section 2.7.1), were heat denatured by boiling, added to the filters and hybridizations allowed to proceed at 65°C overnight. Filters were subjected to three 5 minute washes with 1litre 0.1X SSC, 0.1% SDS at room temperature, followed by two 30 minutes washes with 1litre 0.5X SSC, 0.1% SDS at 65°C. Filters were then blotted dry and exposed to X-ray film as above.

In one series of experiments (section 3.2.2), hybridizations were performed under the conditions at which the  $T_m$  of the cloned genomic fragments had been previously determined (A. Lohe, *pers. comm.*). The filters were pre-hybridized in 3X SSC, 50% formamide, 5X Denhardt's solution and 500ug/ml sonicated, denatured herring sperm DNA for 4 to 8 hours at 25°C ( $T_m - 15^\circ\text{C}$  for the sequence with the lowest  $T_m$  value). Hybridizations were performed in a freshly prepared prehybridization solution, containing only 250ug/ml carrier, at 25°C overnight. Filters were then washed extensively in 3X SSC, 50% formamide at 25°C, dried and exposed to X-ray film as above.

Nitrocellulose filters could be reused several times by eluting bound radioactive DNA with several 60 minute washes with 250ml 40mM NaOH, followed by neutralization in 1.5M NaCl, 0.5M Tris-HCl, pH 7.4, and rinsing in 2X SSC.



### 2.7.3 Northern blot hybridizations

RNA samples that were to be hybridized with double stranded DNA hybridization probes were transferred to nitrocellulose filters as described by Thomas (1980). Gels were placed directly onto the blotting towers and the denatured RNA transferred to nitrocellulose by blotting overnight in 20X SSC. Filters were then baked at 80°C *in vacuo* for 2 to 4 hours. Residual glyoxal was removed from filters by washing in 20mM Tris-HCl, pH 8.0 at 100°C.

Filters were prehybridized overnight at 42°C in 50% formamide (deionized immediately before use by at least three passages through mixed bed resin), 4X SSC, 1X Denhardt's solution, 50mM sodium phosphate buffer, pH 7.0 and 250ug/ml sonicated, denatured herring sperm DNA (Thomas, 1980). Hybridization probes (at least  $3 \times 10^7$  cpm) were denatured by boiling and added directly to the filters. Hybridizations were allowed to proceed for 20 to 24 hours at 42°C. Filters were subjected to four 5 minutes washes with 1 litre 2X SSC, 0.1% SDS, at room temperature, followed by two 15 minute washes with 1 litre 0.1X SSC, 0.1% SDS, at 50°C. Filters were air-dried and exposed to X-ray film as described for Southern blots.

Where RNA samples were to be hybridized to single stranded RNA probes, the RNA was transferred to the nylon-based membrane, Zetaprobe. This was achieved by blotting gels as described above, except that 10X SSC (K. Reed, *pers. comm.*) was used as the transfer medium. Filters were baked and washed as described above.

The prehybridization solution (K. Reed, *pers. comm.*) consisted of 10X Denhardt's solution or 0.5% dried low fat milk powder (R. Saint *pers. comm.*), 5X SSPE, 0.5% SDS, 50% formamide (deionized as above), 500ug/ml sonicated, denatured herring sperm DNA, 500ug/ml carrier yeast RNA and 20ug/ml poly A. The filters were prehybridized for 4 to 8 hours at 60°C. Hybridizations were performed in freshly prepared prehybridization solution containing only

250ug/ml carrier DNA. Hybridization probes were denatured in 50% formamide at 70°C for 5 minutes, added directly to the filters and hybridizations allowed to proceed at 60°C overnight. Filters were then subjected to two 15 minute washes with 1 litre 3X SSC, 1.0% SDS at room temperature, followed by two 30 minute washes with 1 litre 0.1X SSC, 1.0% SDS, each at 65°C. Filters were air-dried and exposed to X-ray film as described previously.

## 2.8 Drosophila Techniques

### 2.8.1 Drosophila culture

All stocks were cultured at room temperature (21 to 23°C), and mutant and deficiency stocks were maintained over balancer chromosomes. Experiments to generate organisms homozygous deficient for defined complementation groups were carried out at room temperature or 25°C, and those to produce homozygous *uncoordinated* mutants at 25°C. Organisms of the required genotype were selected under a dissecting microscope, dried, transferred to plastic vials snap frozen in liquid nitrogen and stored at -90°C.

### 2.8.2 Collection of life stages

Organisms representing each major stage of the *Drosophila* life cycle (Figure 2.1) were collected. Large quantities of embryos were obtained by establishing a population cage containing approximately 20,000 adult flies. It was maintained at room temperature with damp towelling added to increase humidity. Females oviposited onto trays of culture media (11.5cm x 20cm x 2.5cm) for 18 hour periods. Egg laying was stimulated by coating the media with blackcurrant fruit juice syrup and applying strips of yeast paste. The embryos were harvested by gentle brushing of the media and floating in T.E. buffer and they were collected by straining the resultant suspension through fine cheese cloth. Embryos were washed repeatedly with T.E. buffer until free

of media and yeast, and the embryonic chorion removed by washing with 50ml 2.5% sodium hypochlorite. The embryos were then rinsed with T.E. buffer, briefly blotted dry, transferred to glass vials and snap frozen in liquid nitrogen.

Small bottles of media were placed in the population cages and the females permitted to lay for 6 hour periods. The bottles were maintained at room temperature and collections of larvae, pupae and adults subsequently made. Larvae were collected at two developmental stages, the first of which represented a mixture of first and second instar larvae, collected 2 to 3 days after egg laying. The second group of larvae collected were third instar larvae (5 days after egg laying). Larvae were floated from the media using 20% sucrose (Ashburner and Thompson, 1978), recovered by straining through cheese cloth and washed with double distilled water. The larvae were briefly blotted dry, transferred to 10ml plastic tubes and snap frozen.

Pupae were collected 7 to 8 days after egg laying, at approximately the mid-pupal stage. Pupae were gently dislodged from the side of the bottle using a soft brush and washed into a wide, shallow beaker with double distilled water. Under these conditions the floating pupae were skimmed from the water using plastic coated guaze. Pupae were dried and stored as described previously.

Adult flies of mixed sexes were collected and starved for 4 hours. For experiments requiring either male or female adults, flies were immobilized after starvation by chilling on ice and then sorted according to sex. Adults were stored in glass vials as described above.

### 2.8.3 Preparation of salivary gland chromosomes

Salivary gland polytene chromosomes were prepared from organisms carrying X chromosomes from Canton-S and an *uncoordinated* mutant. The organisms were cultured in uncrowded conditions at 18°C, with a yeast suspension added daily.

Chromosomes were prepared by a modification of the method of Harshman (1977). Briefly, a microscope slide was wiped with siliconized paper, a ring of lubricant applied and the ring filled with lacto-aceto-orcein stain (1% orcein, 30% acetic acid, 45% lactic acid). Salivary glands were dissected from late third instar larvae in *Drosophila* Ringer's buffer (0.75% NaCl, 0.035% KCl, 0.021%  $\text{CaCl}_2$ ), the fat bodies removed and the glands transferred to the stain for 6 to 9 minutes. A coverslip was then dropped onto the stained glands, the cells broken and chromosomes spread by gentle rotation of the coverslip. The chromosomes were examined under a compound light microscope and compared with photographs of Lefevre (1976). Suitable chromosome complements were photographed at a 100-fold magnification.

Table 2.1

Bacterial Strains

| <u>Strain</u>                                      | <u>Characteristics</u>   | <u>Reference</u>  |
|--|--|---|
| RR1<br>( <i>E. coli</i> K12 X<br><i>E. coli</i> B) | F <sup>-</sup> , hsd20(r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> )<br>ara-14, proA2, lacY1,<br>galK2, rpsL20 (sm <sup>r</sup> ),<br>xyl-5, mtl-1, supE44, λ <sup>-</sup> | Bolivar <i>et al.</i> ,<br>1977                                 |
| LE 392<br>( <i>E. coli</i> K12)                    | F <sup>-</sup> , hsd514(r <sub>k</sub> <sup>-</sup> , m <sub>k</sub> <sup>-</sup> ),<br>supE44, supF58,<br>lacY1, galK2, galT22<br>metB1, trpR55, λ <sup>-</sup>                         | Derived from<br>strain ED8654<br>Murray <i>et al.</i> ,<br>1977 |

Table 2.2

DNA Vectors

| <u>Vector</u> | <u>Characteristics</u>                              | <u>Reference</u>                |
|---------------|---|---------------------------------|
| pBR322        | Ap <sup>r</sup> , Tc <sup>r</sup>                   | Bolivar <i>et al.</i> ,<br>1977 |
| pBR328        | Ap <sup>r</sup> , Tc <sup>r</sup> , Cm <sup>r</sup> | Soberon <i>et al.</i> ,<br>1980 |
| pACYC184      | Tc <sup>r</sup> , Cm <sup>r</sup>                   | Chang and Cohen,<br>1978        |
| pSP65         | Asymmetric cloning sites<br>in M13 polylinker       | Krieg and Melton,<br>1984       |



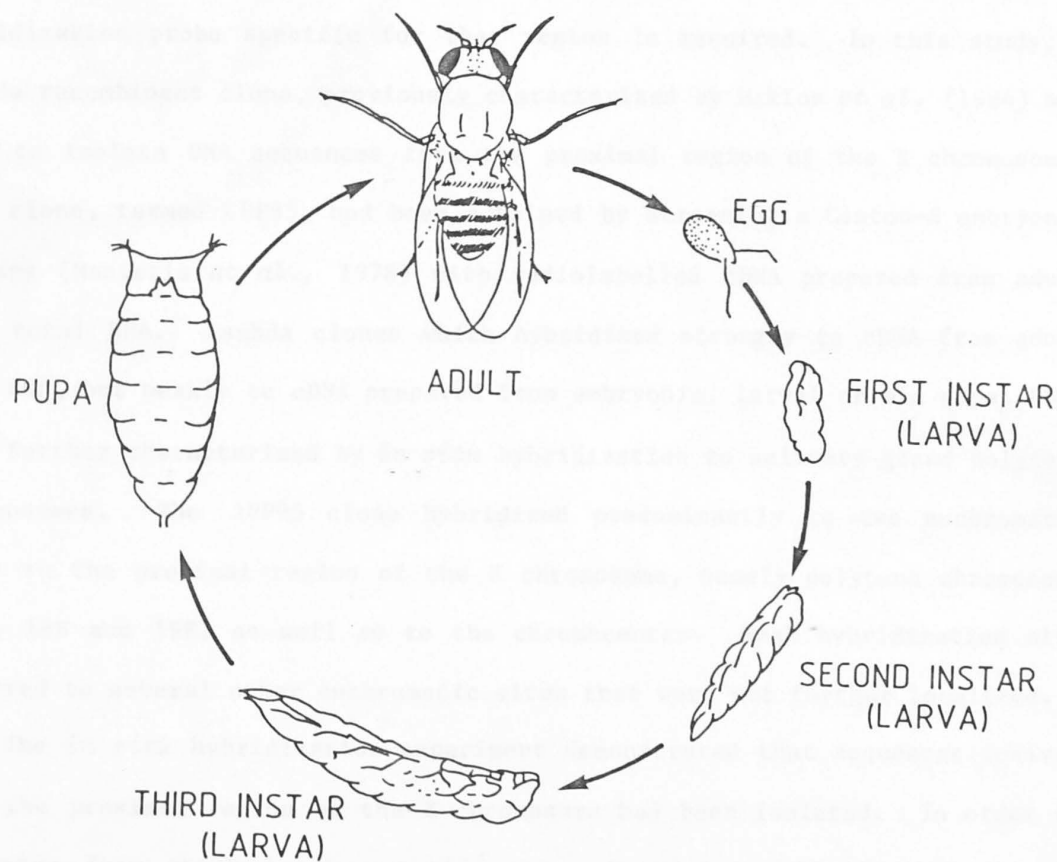
Table 2.3

Drosophila X Chromosomes

| <u>Chromosome</u>                  | <u>Characteristics</u>   | <u>Reference</u>          |
|------------------------------------|--|---------------------------|
| Df(1) B57                          | Df breakpoints: 19E1/E2; 19E8/F1   | Schalet and Lefevre, 1976 |
| Df(1) Q539                         | Df breakpoints: 19E6/E7; 19F6/20A1   | Schalet and Lefevre, 1976 |
| Df(1) S54                          | Df breakpoints: 19E7/E8; 20A1/A2   | Lefevre (unpublished)     |
| y <sup>+</sup> Ymal <sup>106</sup> | Dp breakpoints: 19A3/A4; 20F/Xh  | Schalet and Lefevre, 1976 |
| FM6                                | ln(1) FM6 y <sup>31d</sup> sc <sup>8</sup> dmb   | Lindsley and Grell, 1968  |
| FM7                                | ln(1) FM7 sc <sup>8</sup> + 15D-E; 20A-E + d149,y <sup>31d</sup> sc <sup>8</sup> , w <sup>a</sup> v <sup>0ff</sup> B | Merriam, 1968             |
| Binsn                              | ln(1) sc <sup>SIL</sup> sc <sup>8R</sup> + d1-49   | Lindsley and Grell, 1968  |
| unc <sup>S36</sup>                 | uncoordinated phenotype  | Lefevre (unpublished)     |
| unc <sup>S95</sup>                 | uncoordinated phenotype  | Lefevre (unpublished)     |
| unc <sup>DC803</sup>               | uncoordinated phenotype  | Lefevre (unpublished)     |
| unc <sup>GE230</sup>               | uncoordinated phenotype  | Lefevre (unpublished)     |
| unc <sup>16.3.212</sup>            | uncoordinated phenotype  | Schalet and Lefevre, 1976 |
| unc <sup>W5</sup>                  | uncoordinated phenotype  | Lifschytz and Falk, 1969  |
| unc <sup>27E2</sup>                | uncoordinated phenotype  | Schalet (unpublished)     |
| Canton-S (Caltech)                 | wildtype   | Lindsley and Grell, 1968  |
| M56i (Amherst)                     | wildtype   | Schalet and Lefevre, 1976 |

Figure 2.1: The *Drosophila melanogaster* life cycle.

During the 22-24 hours (at 25°C) of embryonic development, differentiation of the larva occurs. The larva grows primarily by moulting and it passes through three instars and two moults. At each moult the entire cuticle, and specialized cuticular structures, are shed and the structures characteristic of the ensuing instar differentiate. The moulting process affects mainly the body wall, whereas growth of the internal organs is largely independent of the moults and occurs by an increase in cell size rather than by cell multiplication. In contrast, imaginal discs, which are the progenitors of adult structures, grow by cell multiplication. The series of developmental steps by which the insect passes from the larval to the adult stage is called metamorphosis. The formation of the puparium, which is identical to the cuticle of last larval instar, occurs at about 96 hours after egg hatching. The organism then forms a prepupa by separating the epidermis from the puparium and pupates by everting its mouth-parts, wings, halteres and legs. The prepupal cuticle is shed and the pupal cuticle formed. During this metamorphic stage, some larval tissues and organs undergo histolysis and are formed anew. The imaginal discs differentiate to form adult structures, and also form the body wall. The imago emerges as an adult fly at approximately 192 hours egg hatching at 25°C. (adapted from Bodenstein, 1950).



## CHAPTER THREE

ISOLATION AND ANALYSIS OF GENOMIC DNA FROM THE UNCOORDINATED COMPLEMENTATION GROUP

In order to isolate genes from a defined chromosomal position, a hybridization probe specific for that region is required. In this study, a lambda recombinant clone, previously characterized by Miklos *et al.* (1984) was used to isolate DNA sequences from the proximal region of the X chromosome. This clone, termed  $\lambda$ PP95, had been obtained by screening a Canton-S embryonic library (Maniatis *et al.*, 1978) with radiolabelled cDNA prepared from adult head total RNA. Lambda clones which hybridized strongly to cDNA from adult head RNA, but weakly to cDNA prepared from embryonic, larval or mid pupal RNA, were further characterized by *in situ* hybridization to salivary gland polytene chromosomes. The  $\lambda$ PP95 clone hybridized predominantly to two euchromatic sites in the proximal region of the X chromosome, namely polytene chromosome bands 19B and 19E, as well as to the chromocenter. Weak hybridization also occurred to several other euchromatic sites that were not further localized.

The *in situ* hybridization experiment demonstrated that sequences derived from the proximal region of the X chromosome had been isolated. In order to determine from which complementation group in division 19 the genomic DNA contained within  $\lambda$ PP95 originated, Miklos *et al.* (1984) constructed genotypes heterozygous for chromosomes carrying genetically defined rearrangements. Southern blots carrying *EcoRI* cleaved genomic DNA, prepared from the adult heads of either heterozygous deficiency genotypes, or the Canton-S wildtype strain, were hybridized with radiolabelled DNA from a recombinant plasmid containing a low copy number sequence of  $\lambda$ PP95, the *EcoRI* fragment p(PP95)-P28. The amount of DNA tested from each genotype was standardized by

hybridization to a recombinant plasmid (*SacI*) carrying the alcohol dehydrogenase (*Adh*) gene and its flanking sequences (Goldberg, 1980). *Adh* is a unique, autosomal gene in *D. melanogaster*, and therefore present in two copies in all strains examined. Comparison of the intensity of the signals obtained using radiolabelled p(PP95)-P28 as a hybridization probe, ascertained the presence or absence of the genomic DNA contained in p(PP95)-P28 in chromosomes rearranged for different regions of division 19. Quantitation of the results using a laser microdensitometer localized the p(PP95)-P28 *EcoRI* fragment to the *uncoordinated* (*unc*) complementation group, formally placed in the polytene chromosome band 19E8.

### 3.1 Isolation of Genomic DNA from the *uncoordinated* Complementation Group

#### 3.1.1 Isolation of lambda clones by chromosome walking

Genomic DNA was isolated from the *uncoordinated* complementation group using the technique of 'chromosome walking' (Bender *et al.*, 1983(b)). This procedure involves the isolation of a series of overlapping DNA segments from a library of cloned genomic sequences. In this study, sequences were isolated in both directions from the cloned DNA contained within the recombinant bacteriophage  $\lambda$  PP95. DNA prepared from  $\lambda$  PP95 was cleaved with the restriction enzyme *EcoRI* and the resultant fragments subcloned into a plasmid vector (see Table 3.1 for the nomenclature and description of plasmids). DNA from the recombinant plasmids containing low copy number sequences, (p(PP95)-P1 and p(PP95)-P28), were used as hybridization probes for screening a genomic library, which had <sup>been</sup> constructed by cloning randomly sheared Canton-S embryonic DNA into the *EcoRI* site of the vector  $\lambda$  Charon 4 (Maniatis *et al.*, 1978). *EcoRI* fragments from each overlapping recombinant bacteriophage were further subcloned into plasmid vectors (Table 3.1) and low copy number sequences used as subsequent hybridization probes. Using this technique, 60kbp of genomic DNA surrounding  $\lambda$  PP95 was isolated (Figure 3.1).

Despite extensive screening of the Canton-S genomic library (summarized in Table 3.2), no further lambda clones were identified which contained overlapping genomic sequences, resulting in the blockage of the chromosome walk in both directions. In an attempt to overcome the blockage, two further genomic libraries were screened, as strain polymorphisms in the location of repetitive elements have enabled blockages in chromosomal walks at other loci to be overcome, for example, the *bithorax* and *white* loci (Bender *et al.*, 1983(a); Pirrotta *et al.*, 1983). The first library was constructed by cloning randomly sheared DNA from a second *D. melanogaster* wildtype strain, Oregon-R, into the vector,  $\lambda$  Sep 6 (Meyerowitz and Hogness, 1982) and the second library, by cloning a partial *Sau*3A digest of Canton-S adult head DNA into  $\lambda$  EMBL3a (Frischauf *et al.*, 1983). The second library was therefore constructed from the same source of DNA as that used in the genomic Southern blot hybridization analyses.

The Oregon-R  $\lambda$  Sep 6 library was screened with four unique *Eco*RI fragments and the Canton-S  $\lambda$  EMBL3a with five (Table 3.2). However, no further overlapping clones were identified.

### 3.1.2 Determination of the restriction enzyme cleavage maps of lambda clones

The restriction enzyme cleavage map of the cloned DNA, together with the overlapping recombinant clones isolated during the chromosomal walk, are shown in Figure 3.1. Restriction maps were determined by cleaving DNA prepared from each recombinant lambda clone, and the corresponding subclones, with a series of restriction enzymes, including *Eco*RI, *Hind*III, *Pst*I, *Pvu*II, *Bam*HI, *Xba*I and *Kpn*I. After digestion with the enzymes, both singly and in combination, the molecular sizes of the resultant fragments were determined by agarose gel electrophoresis, although fragments less than 400bp were more accurately sized by acrylamide gel electrophoresis. The restriction maps were confirmed by



hybridization of radiolabelled DNA from each subclone to Southern blots carrying restriction fragments from the lambda clones.

Extensive cross-hybridization, summarized in Figure 3.1, was revealed between fragments within lambda clones, and also between non-overlapping fragments of different clones. For example, the two subclones derived from p(PP95)-P9, namely p(PP95)-P9.7 and p(PP95)-P9.8, which cross-hybridize only weakly, showed strong hybridization with p(MH26)-C19, p(GR1)-M8 and p(RJ2)-B3 DNA. In addition, the P9.8 fragment hybridized strongly with DNA of p(MH26)-A22 and p(RJ7)-1-9 (Figure 3.2). Therefore, to determine that the lambda clones overlapped in the manner illustrated in Figure 3.1, more detailed restriction maps were constructed for each of the *EcoRI* fragments that cross-hybridized with p(PP95)-P9.7 or p(PP95)-P9.8 (Figure 3.3)\*. Although the data indicates that the lambda clones RJ7 and RJ2 overlap by 3.0kbp, it is possible that  $\lambda$ RJ7 lies slightly further upstream, as it is proposed that the upstream region contains related repeated sequences.

The mapping data revealed a very unusual structural arrangement of sequences within the 60kbp of cloned DNA. Within each of the lambda clones MH26, GR1, RJ2 and RJ7, there is a unit of at least 6.2kbp, extending from the *EcoRI* fragment of 0.24kbp to that of 0.89kbp (see Figure 3.1), which encompasses six different fragments, none of which cross-hybridize (Figure 3.4). Adjacent to both ends of the unit are *EcoRI* fragments which hybridize strongly to p(PP95)-P9, but each of which varies in length and restriction enzyme cleavage sites. These end fragments are members of a DNA sequence family present at a high frequency in the genome (Section 3.2). Thus, this 6.2kbp unit is repeated at least three times in a tandem array with each unit separated by repeated DNA (Figure 3.4). The repeating unit is also found in the clone  $\lambda$ GM2, which was isolated on the basis of hybridization to p(PP95)-P1. It has not been possible to correlate the genomic DNA carried by this

\* The overlapping regions of  $\lambda$ RJ2 and  $\lambda$ RJ7 have in common the presence or absence of cleavage sites recognised by six different endonucleases, each of which recognizes a six base pair target sequence.

clone with the previously described region, although the presence of the repeating unit suggests that it may lie further upstream from  $\lambda$ RJ7.

### 3.2 Molecular Analysis of Genomic DNA from the *uncoordinated* Complementation Group

In *D.melanogaster*, repeated sequences constitute a significant component of centromeric heterochromatin (section 1.1.2) and may also be associated with intercalary heterochromatin (section 1.1.3). Cytological and genetic data have indicated that subdivision 19E is a site of intercalary heterochromatin (Mukhina *et al.*, 1981; Zhimulev, 1982) and also a preferential site of insertion of certain repeated sequences (Berg *et al.*, 1980; Bingham *et al.*, 1982; Green and Miklos 1983; Belyaeva *et al.*, 1984). It was therefore of interest to assess if the cytological and genetic properties could be correlated with the molecular organization of the *uncoordinated* complementation group.

#### 3.2.1 Genomic representation of DNA sequences from the *uncoordinated* complementation group

The relative genomic abundance of the DNA sequences from the *uncoordinated* complementation group was assessed by several experimental approaches. Firstly, radiolabelled genomic DNA prepared from Canton-S adult heads was hybridized to *Eco*RI cleaved DNA from lambda clones spanning coordinates -16.7 to +23.5 (Figure 3.5). If a fragment from the *uncoordinated* complementation group belongs to a high copy number sequence family, the additive effect of related sequences present at many sites in the genome would be sufficient to produce a hybridization signal. The fragments C19, A22 and P9.8 (coordinates -16.7 to -13.3, -13.3 to -11.7 and -5.2 to 0), which show strong homology to each other, hybridized with the genomic DNA. However,

fragments W8 (coordinates -5.1 to -6.0), and C7, C8 and C5 (coordinates +4 to +16.0) showed only weak hybridization to the genomic DNA.

The reciprocal experiment was also performed, whereby DNA prepared from subcloned fragments was hybridized to Southern blots carrying *EcoRI* cleaved genomic DNA prepared from Canton-S adult heads. Adult heads were used as the source of genomic DNA as the chromosomes from the tissues within the head are considered to be primarily diploid as was shown by comparing the genomic representation of low copy number sequences in DNA prepared from adult heads with that from diploid tissues (Miklos *et al.*, 1984). Each *EcoRI* fragment from coordinates +23.5 to -11.0 was tested, except those less than 500bp. As the two *EcoRI*-*PvuII* fragments from the 0 to -5.1 region have little homology, each was used as a hybridization probe. The 2.75, 2.45 and 1.77kbp fragments of  $\lambda$ PP95 (coordinates 0 to -2.75, -2.75 to -5.1 and -6.0 to -7.8, respectively) hybridized to multiple genomic *EcoRI* bands, whereas the remaining fragments detect only a single, or few, genomic bands (summarized in Figure 3.9 (A)).

In order to localize more precisely the repeated sequence within P9.8 (coordinates 0 to -2.75), the fragment was cleaved at the *HindIII* sites to yield three further fragments of approximately equal size (890, 890 and 970 bp). DNA prepared from each subcloned fragment was hybridized to Southern blots carrying *EcoRI* cleaved DNA from Canton-S adult heads (Figure 3.6). Each fragment hybridized intensely to high molecular weight undigested regions, as well as discrete bands, indicating that each fragment contained highly repeated sequences.

The number of plaques hybridizing to each fragment in library screening experiments was also used to estimate the relative abundance of each fragment in the genome (Table 3.3). Fragments W17, P9.7 and P9.8 (coordinates -7.8 to -6.0, -5.1 to -2.45 and -2.45 to 0) hybridized to 10, 115 and 800 plaques per

genome equivalent, respectively. Each of the remaining fragments detected only 1 to 5 plaques per genome equivalent.

The results of these experiments illustrate that the genomic representation of fragments from this region of the *Drosophila* genome is variable. Several different families of repeated sequences have been identified, that are interspersed with low copy number and unique sequences (summarised in Figure 3.7). Three types of repeated sequences were detected. The first category contains sequences exemplified by the P9.8 and P9.7 fragments (coordinates 0 to -2.75 and -2.75 to -5.1), that have extensive regions of homology and by all three criteria belong to a highly repeated sequence family. Repeated sequences in the second category are probably short, but highly abundant in the genome since fragments containing these sequences hybridized to only the expected band on genomic Southern blots. However, when radiolabelled *Drosophila* genomic DNA was hybridized to the cloned sequences, the additive effect from many genomic sites resulted in a hybridization signal. The *EcoRI* fragments C7, C8 and C5 (coordinates +4.0 to +16.0) all fall into this category. The 1.77kbp fragment, at -6.0 to -7.8, appears to belong to a third category. This fragment hybridized to multiple bands on a Southern blot of genomic DNA, which suggested that it contained repeated sequences, but it was not detected by radiolabelled *Drosophila* DNA. This indicates that this sequence is only weakly homologous to a number of genomic regions.

### 3.2.2 Characterization of high copy number sequences

Several classes of repeated sequences have been characterized in both heterochromatic and euchromatic regions (reviewed in Spradling and Rubin, 1981). The repeated sequences identified in the *uncoordinated* complementation group were tested for cross-hybridization to the elements listed in Table 3.4

to determine if they corresponded to previously characterized repeated elements, and in particular, those correlated with sites of either intercalary or centromeric heterochromatin.

Each of the four major *Drosophila* satellite sequences was tested for its presence in the *uncoordinated* region by hybridizing radio-labelled fragments P9.8, P9.7 and W17 to Southern blots carrying different components of each satellite. No homology was detected (data not shown), nor was there any cross hybridization between the 359bp tandem repeat from satellite III (Carlson and Brutlag, 1977) and  $\lambda$ PP95 DNA (Miklos *et al.*, 1984).

Moderately repetitive sequences occur both at euchromatic sites and in the centromeric heterochromatin. DNA of a number of elements from two classes of middle repetitive DNA, the *copia*-like and clustered scrambled repeat families, were hybridized to Southern blots carrying *Eco*RI restricted DNA from lambda clones spanning coordinates -16.7 to +23.5. Two of these elements, the *copia*-like sequence, *mdg-3* (Ilyin *et al.*, 1980), and the scrambled repetitive sequence, pDML (Wensink *et al.*, 1974), were of particular interest as *in situ* hybridization experiments had localized both sequences to subdivision 19E (Belyaeva *et al.*, 1984). However, none of the elements tested cross-hybridized with genomic DNA isolated from the *uncoordinated* complementation group.

The final class of repeated sequences tested was the type 1 insertion, a sequence that disrupts the continuity of the 28S genes in division 20 (Glover and Hogness, 1977). Figure 3.8 shows that the type 1 insertion sequence hybridizes with *Eco*RI fragments of 3.45 and 1.57kbp from  $\lambda$ MH26 and 5.1kbp from  $\lambda$ PP95. However, the type 1 insert has homology with fragment P9.8 (coordinates 0 to -2.75) but not with the P9.7 fragment (coordinates -2.75 to -5.1) (data not shown). Comparison of the restriction enzyme cleavage maps of the type 1 insert and fragment P9.8 (Figure 3.3 and Dawid *et al.*, 1978)



indicates they have only a small region of homology, between probably less than 1.0kbp.

Sequences adjacent to the tandemly repeated unit found in the *uncoordinated* complementation group (section 3.1.2) with homology to the P9.8 fragment contain type 1 insertion sequences, indicating that this region may have a sequence arrangement similar to that described by Kidd and Glover (1980) and Dawid *et al.* (1981), where type 1 insertion sequences occurred as tandem repeats, often interspersed with non-homologous DNA.

Increased hybridization of the P9.8 fragment to DNA prepared from individuals carrying a duplication of divisions 19 and 20 on the Y chromosome suggested that this sequence family was restricted primarily to this region of the chromosome (Miklos *et al.*, 1984). *In situ* hybridization of type 1 sequences to mitotic chromosomes have demonstrated that those sequences not associated with the rRNA genes are located distal to the nucleolus organizer (Peacock *et al.*, 1981; Appels and Hilliker, 1982). Together, these results suggest that the type 1 sequences are, in fact, located in divisions 19 and 20.

### 3.3 Interstrain Variability

#### 3.3.1 Comparison of the sequences from the *uncoordinated* complementation group in two wildtype strains

The region of the *Drosophila* genome in the vicinity of the *uncoordinated* complementation group is variable in terms of restriction enzyme cleavage sites and the frequency of low and high copy number sequence families, both within, and between, laboratory strains. Miklos *et al.* (1984) noted that the genomic fragments detected by p(PP95)-P1 (coordinates 0 to +4), p(PP95)-W8 (coordinates -5.1 to -5.9) and p(PP95)-P9.8 (coordinates 0 to -2.75) differed between the Canton-S strain and certain laboratory stocks. For example, in



some strains two of the three genomic *EcoRI* fragments homologous to p(PP95)-P1 (i.e. the 4.0 and 1.37kbp fragments) were absent, while in other strains the third fragment (4.2kbp) had a restriction enzyme site polymorphism giving two smaller fragments of 1.9 and 2.3kbp. Hybridization of p(PP95)-W8 to the same laboratory strains as before demonstrated that the copy number of this sequence family can vary at least 14-fold.

A comparison of the wildtype strains, Canton-S and M56i (Amherst), also revealed that this region of the genome varies between strains (see Section 5.1). In this study, each subcloned fragment from coordinates +23.5 to -11.0 was hybridized to Southern blots carrying *EcoRI* and *HindIII* cleaved genomic DNA from both wildtype strains (Figure 3.9). A restriction enzyme cleavage map for the M56i strain was deduced from the hybridization data (Figure 3.10). Comparison of the Canton-S and M56i restriction enzyme maps reveals that the two strains are virtually identical from coordinates +23.5 to +4. The only detectable difference is a slight change in the position of the *HindIII* site in the *EcoRI* B46 fragment (coordinates +16.0 to +18.4) and thus in this region, restriction enzyme site polymorphisms occur at the rate of 0.5 per 10kbp of DNA. However, in the remainder of the region an average of 4.5 polymorphisms per 10kbp were found. The most notable difference occurs at coordinates 0 to +4, where an insertional polymorphism resulted in the presence of two *EcoRI* fragments homologous to p(PP95)-P1 (the 4.0 and 4.2kbp fragments). Many differences were detected within the tandemly repeated 6.2kbp unit described in section 3.1.2. However, it is not possible to assess from this data whether all the repeated units in the tandem array are altered. Clarification of this point would require the isolation of these sequences from a library constructed from M56i genomic DNA.

Some of the restriction enzyme site polymorphisms that occur between the M56i and Canton-S strains, may also occur between individuals of the Canton-S

population. For example, the 1.8kbp *EcoRI* fragment (C20) hybridized to two Canton-S *EcoRI* genomic fragments of 1.8 and 5.2kbp (Figure 3.9(A)). Since the 5.2kbp fragment is also found in M56i (Figure 3.9(A)), it is likely that this fragment was generated in the Canton-S strain by a restriction site polymorphism, rather than the genome containing two cross-hybridizing fragments.

### 3.3.2 Comparison of hybridization patterns of repeated sequences from the *uncoordinated* complementation group in different wildtype strains

Middle repetitive DNA sequences that exhibit interstrain variation are often mobile (Finnegan *et al.*, 1977). Mobile elements are characterized by a high number of copies per genome (20-30) and are present at many different sites which may vary between strains. Because the DNA sequences from the *uncoordinated* complementation group exhibit both inter- and intra-strain variability (see Section 3.3.1), the repeated sequence families were tested for mobility by comparing their hybridization patterns with genomic DNA prepared from Canton-S and two other wildtype strains, F/F<sup>+</sup> and SF1/SF1. The F/F strain is homozygous for the *Adh-F* allele on the second chromosome and was derived from a single pair mating of flies collected in New South Wales, Australia (Lewis and Gibson, 1978). The SF1/SF1 strain was collected from Iowa, U.S.A. and is homozygous for a single wildtype second chromosome (Sampsell, 1977). Fragments P9.8 and W17 (coordinates -2.75 to 0 and -7.8 to -6.0, respectively) were hybridized to Southern blots carrying *EcoRI* restricted DNA from the wildtype strains Canton-S, F/F and SF1/SF1 (Figure 3.11).

Comparison of the hybridization patterns reveal both quantitative and qualitative differences. For example, the intensity of hybridization of both fragments to DNA from the F/F strain is considerably less than that to Canton-

S and SF1/SF1 DNA, although similar amounts of DNA were tested. In addition, the F/F DNA did not contain 3.75 and 2.35kbp *Eco*RI fragments homologous to W17 and contained high molecular weight sequences homologous to P9.8. Hybridization of the P9.8 and W17 repetitive sequences to *Eco*RI cleaved M56i DNA also revealed variability in both the intensity and patterns of hybridization (Figure 3.9(A)).

The repeated sequences varied between strains in both the size of the genomic fragments that contained the sequence, and the relative abundance of each sequence. Although these properties are characteristic of evolutionarily mobile sequences, the mechanism responsible for the strain variability of these sequences cannot be determined. Sequences may be dispersed within the genome by mechanisms including transposition of mobile elements (Strobel *et al.*, 1979) or unequal sister chromatid exchange (Smith, 1976). The two possibilities would be discriminated from an analysis of each repeated sequence to determine if its structural features were characteristic of mobile elements (Rubin *et al.*, 1980).

### 3.4 Correlation of Genetic and Molecular Maps

Mapping experiments of Miklos *et al.* (1984) localized the first cloned entry point into divisions 19 and 20,  $\lambda$  PP95, to the *uncoordinated* complementation group, formally placed in polytene chromosome band 19E8. This was achieved using genotypes heterozygous for chromosomal rearrangements whose breakpoints had been defined genetically. Further mapping experiments were performed as part of this study, using genotypes heterozygous for overlapping deficiencies to produce individuals homozygous deficient for only the *uncoordinated* complementation group (see Section 4.5). This series of experiments had a two-fold purpose. Firstly, it was necessary to ensure that all genomic fragments isolated during the chromosome walk had originated from

the *uncoordinated* complementation group. Because several of the fragments were shown to be members of repeated sequence families, the possibility existed that DNA was isolated from other regions of the genome. Secondly, if a genomic fragment were found to cross the breakpoint of a chromosomal rearrangement, it would be possible to orient the cloned DNA relative to the chromosome.

Genomic DNA was isolated from late pupae of Canton-S, B57/FM6, B57/Q539 and B57/S54 genotypes. B57/Q539 individuals are homozygous deficient for DNA sequences in two complementation groups, *varied outspread* (*vao*) and *uncoordinated*, B57/S54 individuals only lack the DNA sequences from *uncoordinated* and B57/FM6 individuals are heterozygous for the B57 deficiency (Figure 4.9). Southern blots carrying *Eco*RI cleaved genomic DNA from late pupae of each genotype, were successively hybridized with DNA from subcloned fragments from coordinates +23.5 to -11.0. Figure 3.12 demonstrates that each fragment tested was contained within the deficiency encompassing only the *uncoordinated* complementation group.

These data confirmed that all cloned sequences originated from within the *uncoordinated* complementation group, and implies that the size of this complementation group is at least 60kbp. However, as the cloned sequences do not extend across the breakpoints of either the B57 or S54 chromosomal rearrangements, it was not possible to orient the molecular map relative to the chromosome.

### 3.5 Conclusions

This study resulted in the isolation of 60kbp of genomic DNA from a defined euchromatin/heterochromatin transition zone. The molecular and genetic maps were correlated using genetically defined chromosomal deficiencies for varying sections of division 19. These experiments mapped

the cloned genomic DNA to the *uncoordinated* complementation group, and demonstrated that the entire cloned region was contained within this complementation group.

Detailed restriction mapping indicated that the cloned DNA can be divided into two distinct regions. The structural organization of the first region, extending from coordinates +23.5 to +4.0, appears to be unique and is well conserved between two wildtype strains, Canton-S and M56i. This organization is in marked contrast to that of the remainder of the cloned sequences, which is characterized by units repeated in tandem arrays but separated by a member of a high copy number sequence family. At least part of the type 1 insertion sequence of ribosomal genes is a component of the highly repeated sequence, providing further evidence that type 1 insertion sequences occur outside the nucleolus organizer and chromocenter, and demonstrating directly that such sequences may be found adjacent to low copy number coding sequences. This region is also subject to extensive inter- and intra-strain differences, and comparison of the three repeated sequence families in several wildtype strains suggesting variability in both copy number and the location between strains.

Table 3.1

## Description and Nomenclature of Plasmid Subclones

| <u>Clone</u> | <u>Vector</u> | <u>Cloning<br/>Sites (a)</u> | <u>Insert Size<br/>(b) (kbp)</u> | <u>Origin</u>  |
|--------------|---------------|------------------------------|----------------------------------|----------------|
| p(RJ43)-B46  | pACYC184      | R-R                          | 2.37                             | $\lambda$ RJ43 |
| p(RJ43)-H2   | pACYC184      | R-R                          | 2.20                             | $\lambda$ RJ43 |
| p(RJ43)-B26  | pACYC184      | R-R                          | 1.86                             | $\lambda$ RJ43 |
| p(RJ43)-B29  | pACYC184      | R-R                          | 1.12                             | $\lambda$ RJ43 |
| p(RJ43)-B38  | pACYC184      | R-R                          | 0.94                             | $\lambda$ RJ43 |
| p(RJ32)-D7   | pACYC184      | R-R                          | 5.07                             | $\lambda$ RJ32 |
| p(RJ32)-D3   | pACYC184      | R-R                          | 3.42                             | $\lambda$ RJ32 |
| p(RJ32)-E10  | pACYC184      | R-R                          | 3.20                             | $\lambda$ RJ32 |
| p(RJ32)-B12  | pACYC184      | R-R                          | 1.99                             | $\lambda$ RJ32 |
| p(D3/65-1)   | pSP65         | R-R                          | 3.42                             | p(RJ32)-D3     |
| p(D3/65-2)   | pSP65         | R-R                          | 3.42                             | P(RJ32)-D3     |
| p(GM23)-C7   | pACYC184      | R-R                          | 5.4                              | $\lambda$ GM23 |
| p(GM23)-C8   | pACYC184      | R-R                          | 3.32                             | $\lambda$ GM23 |
| p(GM23)-C5   | pACYC184      | R-R                          | 3.15                             | $\lambda$ GM23 |
| p(GM23)-C1   | pACYC184      | R-R                          | 2.62                             | $\lambda$ GM23 |
| p(GM23)-A10  | pACYC184      | R-R                          | 0.53                             | $\lambda$ GM23 |
| p(C7/65-1)   | pSP65         | R-R                          | 5.4                              | p(GM23)-C7     |
| p(C7/65-2)   | pSP65         | R-R                          | 5.4                              | p(GM23)-C7     |
| p(PP95)-P9   | pBR328        | R-R                          | 5.13                             | $\lambda$ PP95 |
| p(PP95)-P1   | pBR328        | R-R                          | 3.93                             | $\lambda$ PP95 |
| p(PP95)-P28  | pBR328        | R-R                          | 2.62                             | $\lambda$ PP95 |
| p(PP95)-W17  | pACYC184      | R-R                          | 1.69                             | $\lambda$ PP95 |
| p(PP95)-W8   | pACYC184      | R-R                          | 0.89                             | $\lambda$ PP95 |
| p(PP95)-P9.8 | pBR328        | R-P                          | 2.75                             | p(PP95)-P9     |
| p(PP95)-P9.7 | pBR328        | R-P                          | 2.45                             | p(PP95)-P9     |
| p(P9)-D14    | pBR322        | R-H                          | 0.97                             | p(PP95)-P9.8   |
| p(P9)-A1     | pBR328        | H-H                          | 0.89                             | p(PP95)-P9.8   |
| p(P9)-E6     | pBR328        | P-H                          | 0.89                             | p(PP95)-P9.8   |
| p(P1/65-1)   | pSP65         | R-R                          | 3.93                             | p(PP95)-P1     |
| p(P1/65-2)   | pSP65         | R-R                          | 3.93                             | p(PP95)-P1     |
| p(GR27)-S11  | pACYC184      | R-R                          | 5.33                             | $\lambda$ GR27 |
| p(MH26)-C19  | pACYC184      | R-R                          | 3.45                             | $\lambda$ MH26 |



|              |          |     |      |                |
|--------------|----------|-----|------|----------------|
| p(MH26) -A28 | pACYC184 | R-R | 2.43 | $\lambda$ MH26 |
| p(MH26) -C20 | pACYC184 | R-R | 1.77 | $\lambda$ MH26 |
| p(MH26) -C31 | pACYC184 | R-R | 1.72 | $\lambda$ MH26 |
| p(MH26) -A22 | pACYC184 | R-R | 1.57 | $\lambda$ MH26 |
| p(MH26) -C9  | pACYC184 | R-R | 1.37 | $\lambda$ MH26 |
| p(MH26) -C34 | pACYC184 | R-R | 0.89 | $\lambda$ MH26 |
| p(MH26) -A32 | pACYC184 | R-R | 0.28 | $\lambda$ MH26 |
| p(MH26) -A25 | pACYC184 | R-R | 0.24 | $\lambda$ MH26 |
| p(GR1) -M11  | pACYC184 | R-R | 3.75 | $\lambda$ GR1  |
| p(GR1) -M8   | pACYC184 | R-R | 3.50 | $\lambda$ GR1  |
| p(GR1) -M16  | pACYC184 | R-R | 1.48 | $\lambda$ GR1  |
| p(RJ2) -B3   | pACYC184 | R-R | 4.78 | $\lambda$ RJ2  |
| p(RJ2) -B21  | pACYC184 | R-R | 1.77 | $\lambda$ RJ2  |
| p(RJ2) -1-5  | pACYC184 | R-R | 1.66 | $\lambda$ RJ2  |
| p(RJ2) -B11  | pACYC184 | R-R | 1.34 | $\lambda$ RJ2  |
| p(RJ2) -B9   | pACYC184 | R-R | 0.78 | $\lambda$ RJ2  |
| p(RJ2) -B17  | pACYC184 | R-R | 0.23 | $\lambda$ RJ2  |
| p(RJ7) -1-22 | pACYC184 | R-R | 2.41 | $\lambda$ RJ7  |
| p(RJ7) -1-9  | pACYC184 | R-R | 1.35 | $\lambda$ RJ7  |
| p(GM2) -X25  | pACYC184 | R-R | 4.30 | $\lambda$ GM2  |
| p(GM2) -X24  | pACYC184 | R-R | 4.15 | $\lambda$ GM2  |
| p(GM2) -W9   | pACYC184 | R-R | 1.80 | $\lambda$ GM2  |
| p(GM2) -X18  | pACYC184 | R-R | 1.71 | $\lambda$ GM2  |
| p(GM2) -X20  | pACYC184 | R-R | 1.40 | $\lambda$ GM2  |
| p(GM2) -X16  | pACYC184 | R-R | 0.89 | $\lambda$ GM2  |
| p(GM2) -X27  | pACYC184 | R-R | 0.27 | $\lambda$ GM2  |
| p(GM2) -X23  | pACYC184 | R-R | 0.23 | $\lambda$ GM2  |

(a) R - *Eco*R1; H - *Hind*III; P - *Pvu*II.

(b) Molecular sizes estimated by averaging 2 to 6 separate determinations

Table 3.2

Summary of Genomic Library Screens

| <u>Genomic Library<br/>Tested</u> | <u>Plasmid<sup>(a)</sup></u> | <u>Lambda clones<sup>(b)</sup></u> |
|-----------------------------------|------------------------------|------------------------------------|
| <u>λ Charon 4 Canton-S</u>        | p(PP95)-P28                  | λ GM23                             |
|                                   | p(PP95)-P1                   | λ GM2, RJ17                        |
|                                   | p(GM2)-2                     | λ RJ2                              |
|                                   |                              | λ RJ7                              |
|                                   | p(GM23)-C8                   | λ RJ32                             |
|                                   | p(RJ32)-B12                  | λ RJ43                             |
|                                   | p(RJ43)-B29                  | nil                                |
|                                   | p(RJ43)-B26                  | nil                                |
|                                   | p(PP95)-W8                   | nil                                |
|                                   | p(GM2)-X27                   | nil                                |
|                                   | p(PP95)-P9.8                 | λ GR1                              |
|                                   |                              | λ GR27                             |
|                                   | p(PP95)-P9.7                 | λ MH26                             |
|                                   | p(RJ43)-B26                  | nil                                |
| <u>λ Sep 6 Oregon-R</u>           | p(PP95)-P1                   | nil                                |
|                                   | p(PP95)-P28                  | nil                                |
|                                   | p(GM2)-W9                    | nil                                |
|                                   | p(PP95)-P1                   | nil                                |
| <u>λ EMBL 3a Canton-S</u>         | p(GM23)-C7                   | nil                                |
|                                   | p(RJ43)-B26                  | nil                                |
|                                   | p(PP95)-W8                   | nil                                |
|                                   | p(RJ43)-B46                  | nil                                |
|                                   |                              | nil                                |

(a) Plasmid subclone used as hybridization probe in library screening

(b) Lambda clones resulting from library screening which extended 'chromosomal walk' (see Figure 3.1)

Table 3.3

Estimation of Genomic Representation of DNA sequences from the  
uncoordinated gene.

| <u>Coordinates of Sequence</u> <sup>(a)</sup> | <u>Copy Number/Genome Equivalent</u> <sup>(b)</sup> |
|---|---|
| -11.0 - -9.2                                  | 3.4 <sup>+</sup>                                    |
| -9.2 - -7.8                                   | 4.6 <sup>+*#</sup>                                  |
| -7.8 - -6.0                                   | 10 <sup>+</sup>                                     |
| -6.0 - -5.1                                   | 2.3 <sup>+*</sup>                                   |
| -5.1 - -2.75                                  | 115 <sup>+</sup>                                    |
| -2.75 - 0                                     | 800 <sup>+</sup>                                    |
| 0 - +4  | 4.6 <sup>+*#</sup>                                  |
| +4 - +9.4                                     | 1.3 <sup>+#</sup>                                   |
| +9.4 - +12.8                                  | 0.8 <sup>+</sup>                                    |
| +12.8 - +16.0                                 | N.T.  |
| +16.0 - +18.4                                 | 2.2 <sup>+</sup>                                    |
| +18.4 - +20.3                                 | N.T.  |
| +20.3 - +22.5                                 | 1.0 <sup>+</sup>                                    |
| +22.5 - +23.6                                 | 1.75 <sup>+</sup>                                   |

(a) Refer to Figure 3.1

(b) Represents the mean of 1 to 3 estimates

(+) Estimate from Charon 4 Canton-S Library Screen

(#) Estimate from EMBL3a Canton-Library Screen

(\*) Estimate from Sep 6 Oregon-R Library Screen

N.T. not tested.

Table 3.4

Repeated Elements tested against DNA sequences from the  
uncoordinated complementation group









| <u>Category of repeated<br/>sequence</u> | <u>Element</u>               | <u>Recombinant<br/>Plasmid*</u>                    |
|--|------------------------------|--|
| Highly Repeated                          | Satellite I<br>(1.672g/cc)   | 1.672-2B <sup>(a)</sup>                            |
|  |                              | 1.672-349 <sup>(a)</sup>                           |
|  |                              | 1.672-453 <sup>(a)</sup>                           |
|  |                              | 1.672-563 <sup>(a)</sup>                           |
|  |                              | 1.672-1 <sup>(a)</sup>                             |
|  |                              | 1.672-38 <sup>(a)</sup>                            |
|  | Satellite II<br>(1.686g/cc)  | 1.686-171 <sup>(a)</sup>                           |
|  | Satellite III<br>(1.688g/cc) | 1.688-198 <sup>(a)</sup>                           |
|  |                              | pDm688.23 <sup>(b)</sup><br>(359bp repeat)         |
|  | Satellite IV<br>(1.705g/cc)  | 1.705-34 <sup>(a)</sup><br>1.705-42 <sup>(a)</sup> |
| Dispersed Moderately<br>Repetitive       |                              |  |
| a) copia-like                            | copia                        | cDm5004 <sup>(c)</sup>                             |
|  | mdg3                         | Dm-86 <sup>(d)</sup>                               |
|  | 412                          | cDm2042 <sup>(e)</sup>                             |
|  | 297                          | cDm4006 <sup>(e)</sup>                             |
|  | Hobo                         | Sgs-4 Isolate <sup>(f)</sup>                       |
| b) scrambled clustered                   | pDml                         | pDml <sup>(g)</sup>                                |
| rDNA Type 1 insert                       | HindIII/Bst 1 4.5Kb          | pC2 <sup>(h)</sup>                                 |

- a) A. Lohe (unpublished)  
 b) Carlson and Brutlag, 1977  
 c) Dunsmuir *et al.*, 1980  
 d) Ilyin *et al.*, 1980  
 e) G.M. Rubin and K. O'Hare (unpublished)  
 f) McGinnis *et al.*, 1983  
 g) Wensink *et al.*, 1974  
 h) Kidd and Glover, 1980

\* The recombinant plasmids were kindly supplied by D. Finnegan (pDm688.23, Dm-86 and pDml), P. Dunsmuir (cDm5004, cDm2042 and cDm4006), S. Beckendorf (Hobo:Sgs-4 Isolate) and R. Appels (pC2).

Figure 3.1: The region cloned from the *uncoordinated* complementation group.

a) Coordinates in units of kbp.

b) Restriction enzyme cleavage map of the cloned DNA, divided into *EcoRI* fragments. The cleavage sites of the restriction enzymes *KpnI*, (K), *XbaI* (X), *PvuII* (p), *PstI* (T), *Hind III* (H), *BamHI* (B) and *SalI* (S) are indicated. Regions with homology to the fragments P1, ; P9.8, ; P9.7, ; W8, ; W17, ; C20, ; A25, ; A32,  are shown. Unshaded regions are unique amongst the cloned sequences.

c) Overlapping recombinant bacteriophage clones divided into *EcoRI* fragments. Fragments used extensively are named.

d) Lambda clone GM2 divided into *EcoRI* fragments and indicating the regions with homology to the fragments in (b).

Group.  
*EcoRI*  
*al* (X),  
 icated.  
 8, ;  
 ons are  
*EcoRI*  
 regions

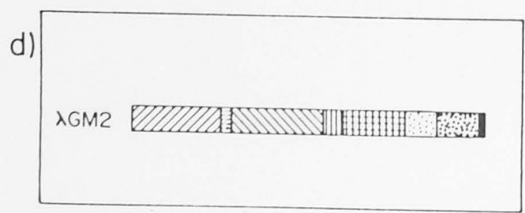
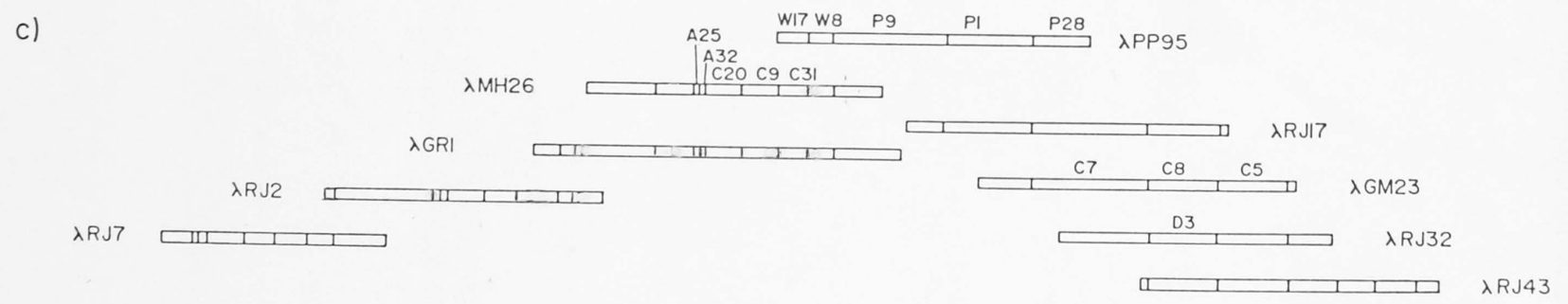
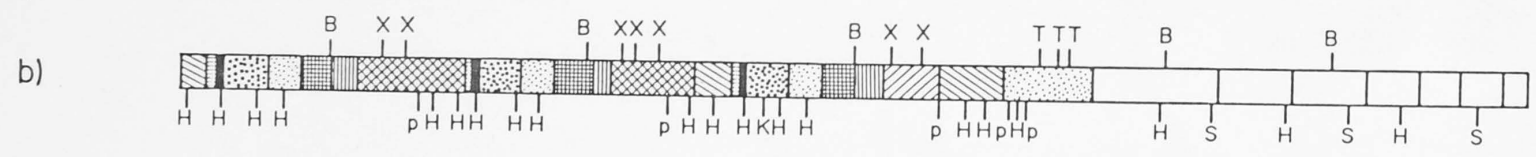
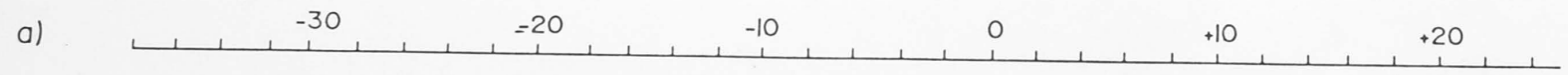




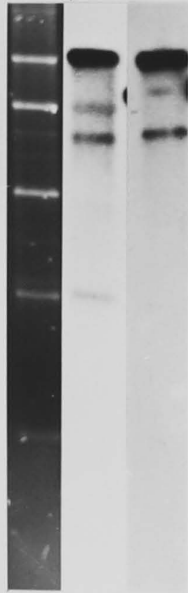
Figure 3.2: Hybridization of p(PP95)-P9.7 and p(PP95)-P9.8 DNA to recombinant bacteriophage clones spanning coordinates -36.6 to +6.6.

DNA from each lambda clone was digested with *Eco*R1, transferred to nitrocellulose after gel electrophoresis and hybridized to radiolabelled DNA of plasmid subclones. Lane 1, the pattern of ethidium bromide staining; lane 2, hybridization to fragment P9.7; lane 3, hybridization to fragment P9.8. Panels a), b), c), d), and e) refer to lambda clones PP95, MH26, GR1, RJ2 and RJ7, respectively. Molecular size markers are indicated in kbp.

binant  
ed to  
ad DNA  
; Lane  
P9.8.  
J2 and

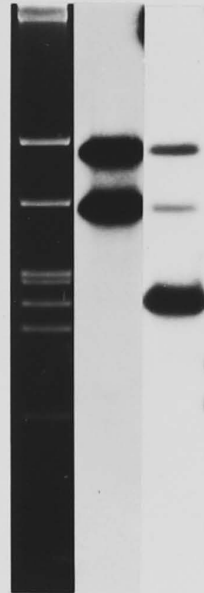
a)

1 2 3



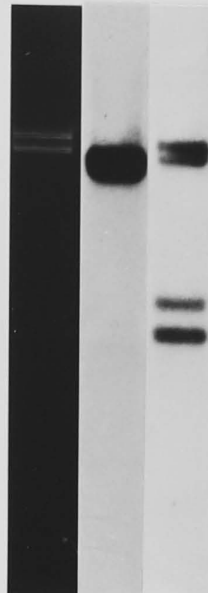
b)

1 2 3



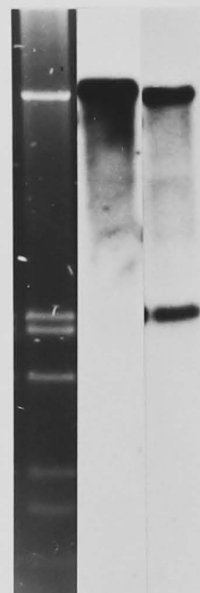
c)

1 2 3



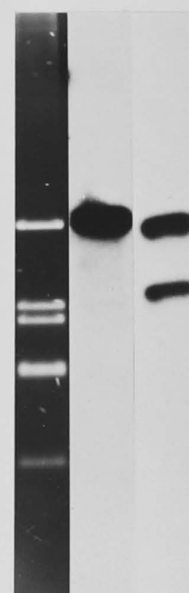
d)

1 2 3



e)




1 2 3



— 6.7  
— 4.3  
  
— 2.3  
— 2.0

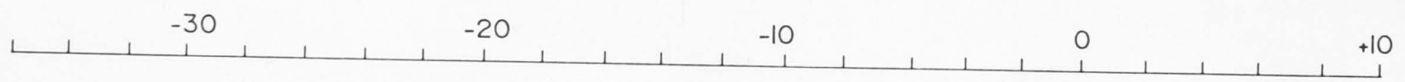
Figure 3.3: Restriction enzyme cleavage maps of recombinant bacteriophage clones spanning coordinates -36.6 to +6.6.

a) Coordinates in units of kbp.

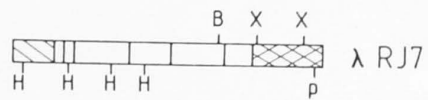
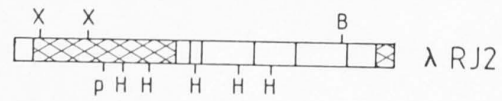
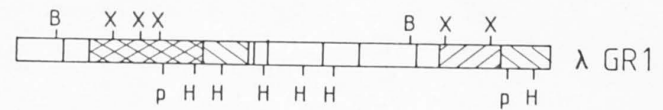
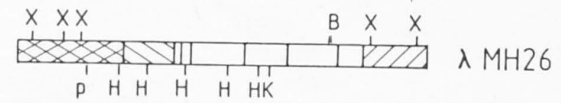
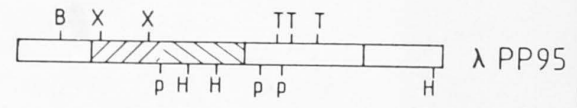
b) Overlapping recombinant lambda clones showing detailed restriction enzyme cleavage map of regions with homology to fragments P9.7 , P9.8. , or both . Each division represents an *EcoRI* fragment. The following restriction enzyme sites are indicated: *EcoRI* (R), *XbaI* (X), *PvuII* (p), *PstI* (T), *HindIII* (H), *KpnI* (K) and *BamHI* (B).

c) Restriction enzyme cleavage map of the recombinant lambda clone GM2.

a)



b)



c)

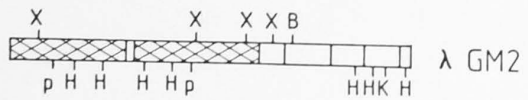






Figure 3.4: Summary of the locations of the 6.2kbp repeated unit.

a) Coordinates in units of kbp.

b) Cloned DNA divided into *Eco*RI fragments. The *Pvu*II (p) cleavage site is also indicated. Regions with homolgy to the repeated unit  and to fragments P9.7  , P9.8  or both  are indicated.

c) Regions with homology to the repeated unit and to fragments P9.7 and/or P9.8 in the recombinant lambda clone GM2.

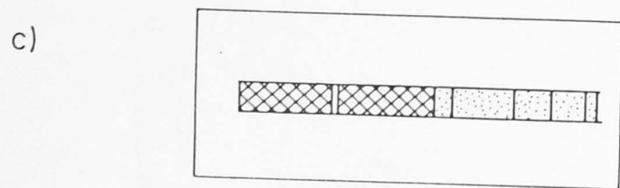
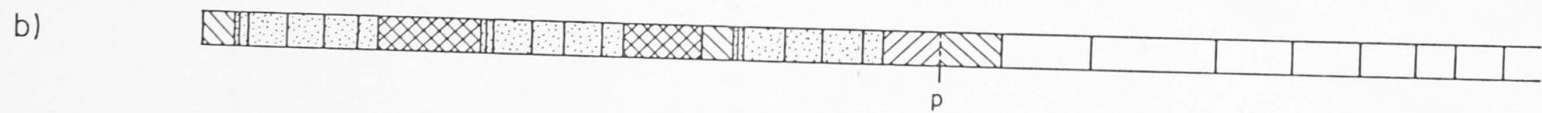
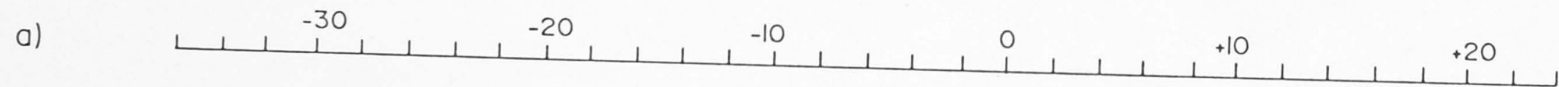




Figure 3.5: Identification of high copy number sequences.

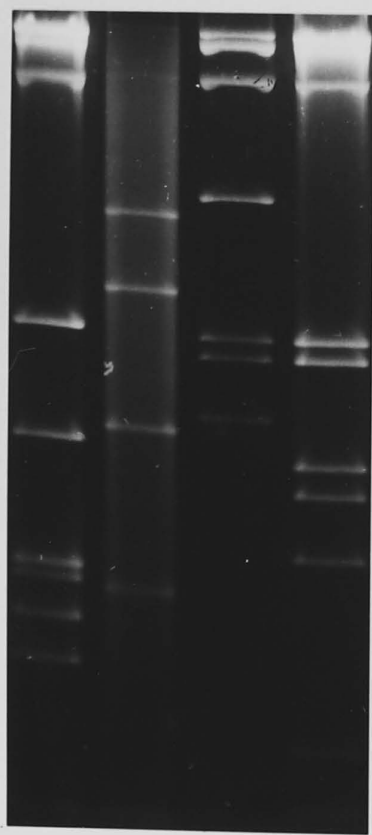
DNA prepared from recombinant bacteriophage clones spanning coordinates -16.7 to +23.5 was cleaved with *EcoRI* and transferred to nitrocellulose after gel electrophoresis.

a) Ethidium bromide staining pattern.

b) Hybridization to radiolabelled total genomic DNA prepared from Canton-S female adult heads. Lanes (1) and (2) were exposed for 3 hours, lanes (3) and (4) for 3 days. The lambda clones tested were: lane 1, MH26; lane 2, PP95; lane 3, GM23; lane 4, RJ43. Molecular size markers are indicated in kbp.

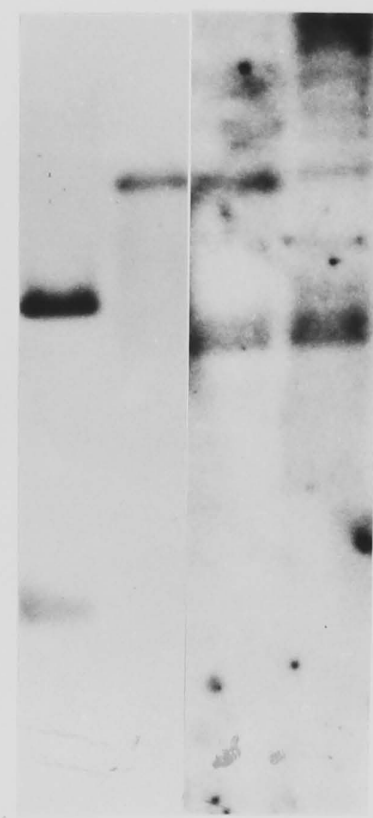
a)

1 2 3 4



b)

1 2 3 4



— 23.7  
— 9.5  
— 6.7  
— 4.3  
  
— 2.3  
— 2.0

Figure 3.6: Genomic representation of subfragments of p(PP95)-P9.8.

Approximately 1 µg genomic DNA extracted from Canton-S female adult heads (lane 1) or male adult heads (lane 2) was cleaved with *Eco*RI and fractionated by gel electrophoresis.

a) The ethidium bromide staining pattern.

b), c) and d) After transfer to nitrocellulose, the genomic DNA was hybridized to radiolabelled DNA from p(P9.8)-A1, p(P9.8)-D14 or p(P9.8)-E6, respectively. Molecular size markers are indicated in kbp.

a)

1 2



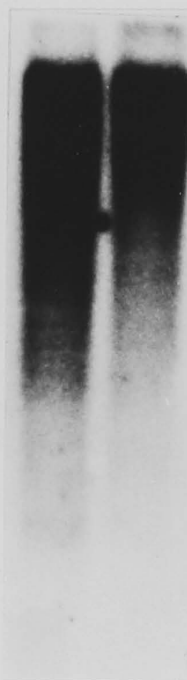
b)

1 2



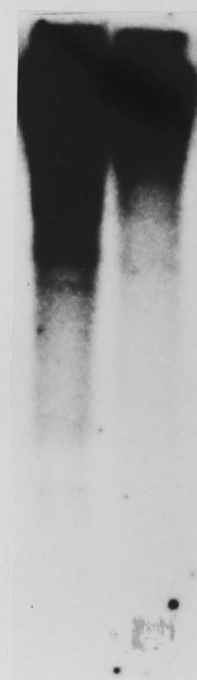
c)

1 2



d)





1 2



— 23.7  
— 9.5  
— 6.7  
— 4.3  
  
— 2.3  
— 2.0

Figure 3.7: Summary of the relative abundance within the genome of fragments from the *uncoordinated* complementation group.

a) Coordinates in units of kbp.

b) Cloned DNA divided into *Eco*RI fragments with the cleavage site of the restriction enzyme *Pvu*II (p) indicated. Sequences are shown which are highly repeated  , show strong homology to multiple genomic sequences  , contain short, highly abundant repeated sequences  , or are weakly homologous to multiple sequences  . The unshaded regions are unique or low copy number.

nts  
the  
hly  
,  
kly  
low

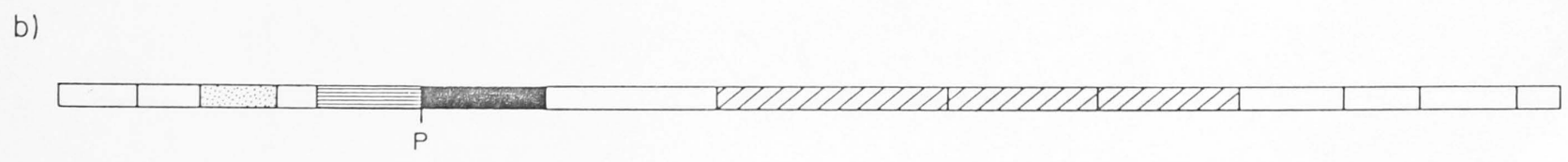
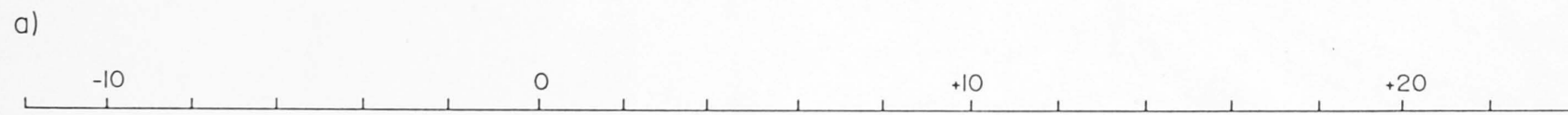




Figure 3.8 Hybridization between the type 1 insert and sequences from the *uncoordinated* complementation group.

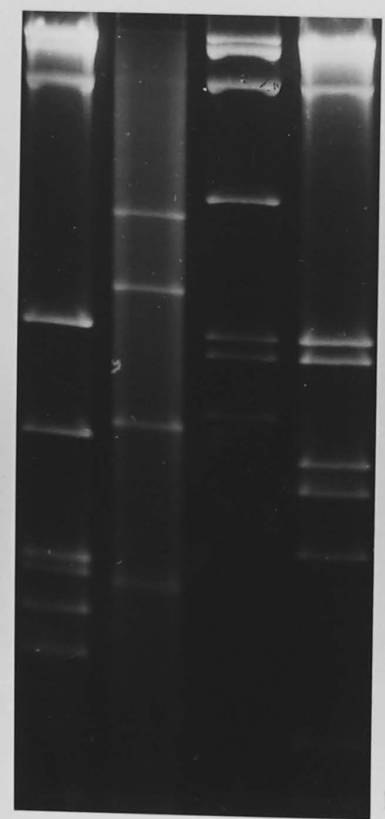
DNA prepared from the recombinant bacteriophage clones spanning coordinates -16.7 to +23.5 was cleaved with *EcoRI* and transferred to nitrocellulose after gel electrophoresis.

a) Ethidium bromide staining pattern.

b) Hybridization with radiolabelled DNA prepared from the type 1 insertion sequence contained within the recombinant plasmid pC2 (see Table 3.4). The lambda clones tested were: lane 1, MH26; lane 2, PP95; lane 3, GM23; lane 4, RJ43. Molecular size markers are indicated in kbp.

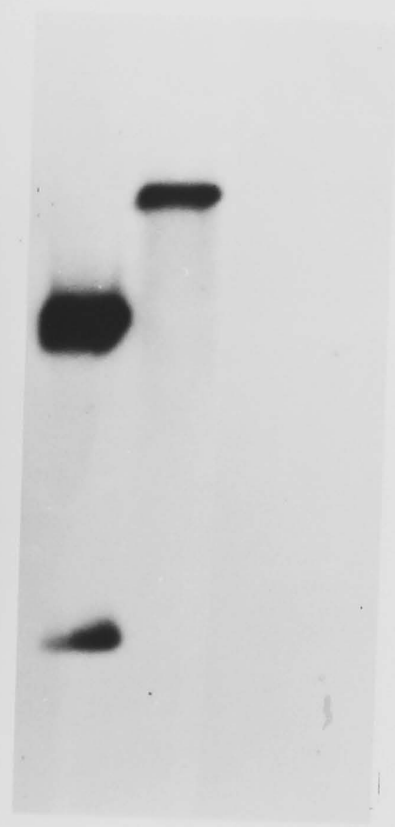
a)

1 2 3 4



b)

1 2 3 4



— 23.7  
— 9.5  
— 6.7  
— 4.3  
  
— 2.3  
— 2.0

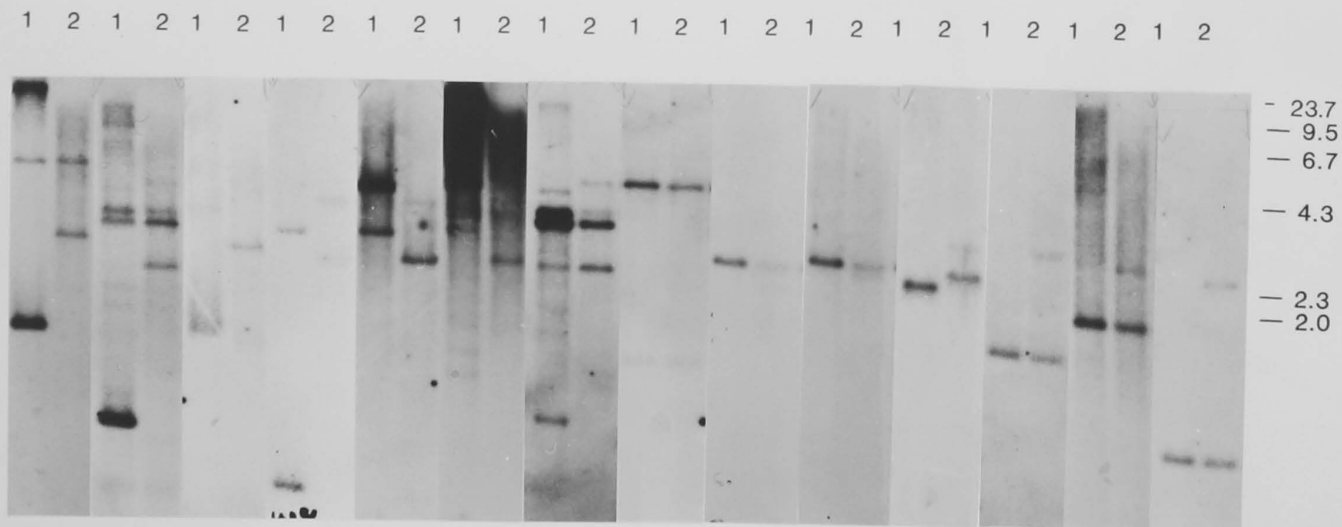
Figure 3.9(A): Genomic representation of the subcloned fragments from the *uncoordinated* complementation group in the wildtype strains, Canton-S and M56i.

a) Approximately 1  $\mu$ g genomic DNA extracted from Canton-S female adult heads (lane 1) or M56i female whole pupae (lane 2) was cleaved with the restriction enzyme *Eco*R1 and separated by gel electrophoresis. After transfer to nitrocellulose, the digested DNA was hybridized to the fragments indicated in (b). The results for the fragments derived from  $\lambda$ PP95 have been described previously (Miklos *et al.*, 1984).

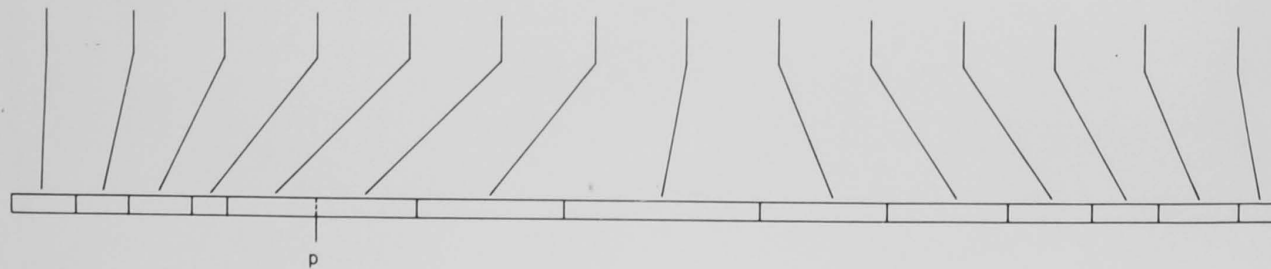
b) Genomic DNA divided into *Eco*R1 fragments, illustrating the fragments used as hybridization probes. The *Pvu*II (p) cleavage site is also indicated.

c) Coordinates in units of kbp.

a)



b)



c)



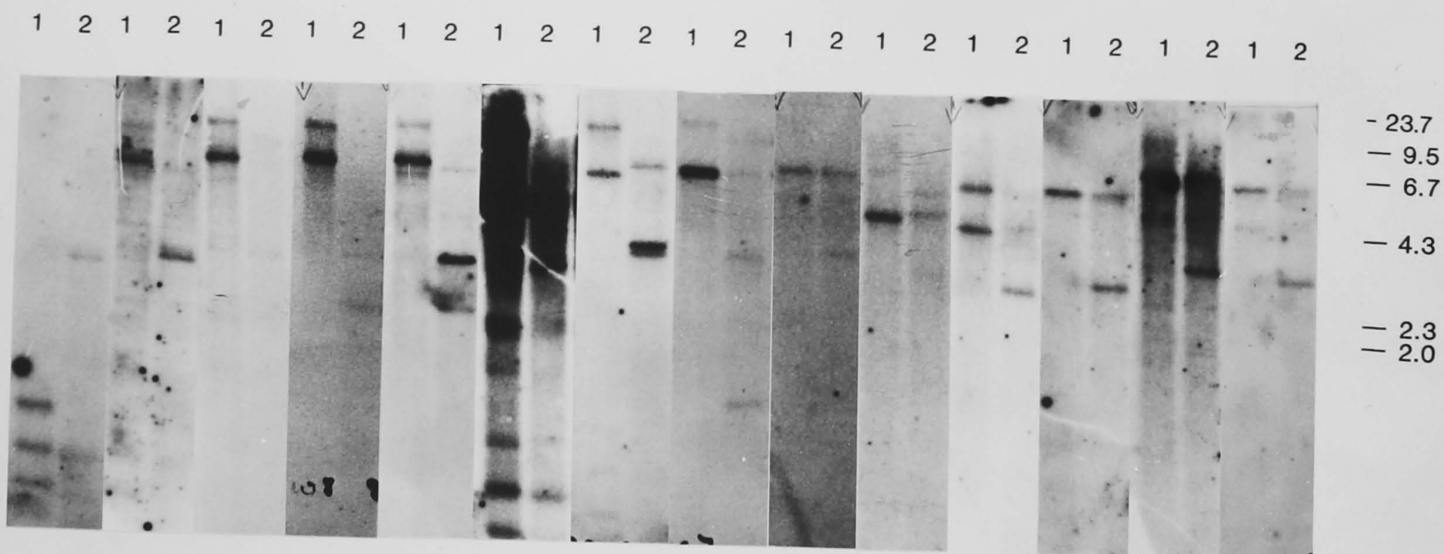
Figure 3.9(B): Genomic representation of the subcloned fragments from the *uncoordinated* complementation group in the wildtype strains, Canton-S and M56i.

a) Approximately 1  $\mu$ g genomic DNA extracted from Canton-S female adult heads (lane 1) or M56i female whole pupae (lane 2) was digested with the restriction enzyme *Hind*III and separated by gel electrophoresis. After transfer to nitrocellulose, the restricted DNA was hybridized to the fragments indicated in (b).

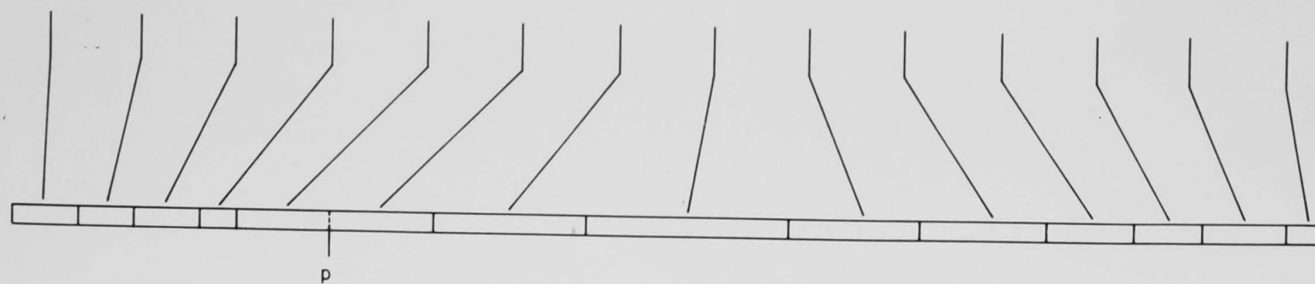
b) Genomic DNA divided into *Eco*RI fragments and illustrating the fragments used as hybridization probes. The *Pvu*II (p) cleavage site is also indicated.

c) Coordinates in units of kbp.

a)



b)



c)

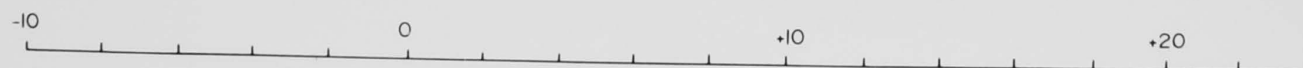




Figure 3.10: Comparison of the restriction enzyme cleavage maps of the cloned region in two wildtype strains, Canton-S and M56i.

a) Coordinates in units of kbp.

b) and c) Restriction enzyme cleavage maps of the strains Canton-S and M56i, respectively. The DNA is divided into *EcoRI* fragments, with the cleavage sites of the restriction enzymes *HindIII* (H) and *PvuII* (p) indicated. An additional *EcoRI* fragment of 4.2kbp is present in the M56i genome at the Canton-S 0 coordinate.

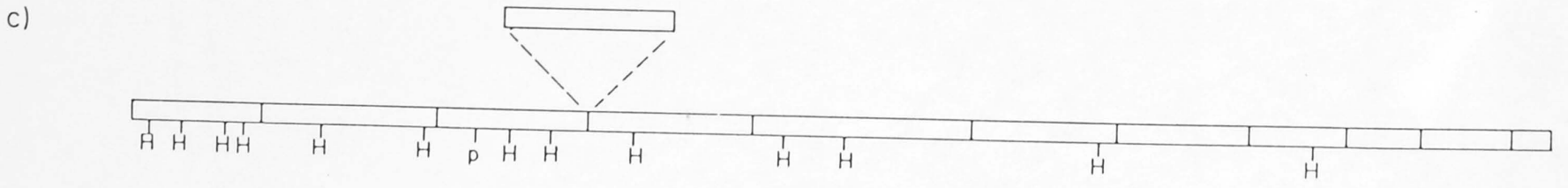
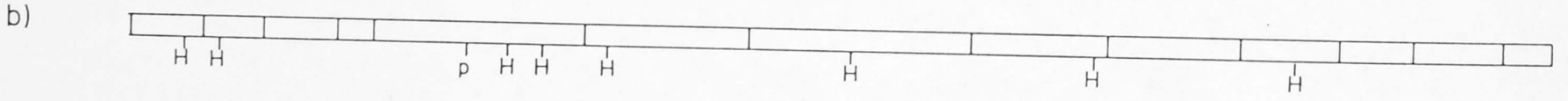


Figure 3.11: Genomic representation of P9.8 and W17 fragments in different wildtype strains.

Approximately 1 µg genomic DNA, extracted from female adult heads of Canton-S (lane 1), F/F (lane 2) and SF1/SF1 (lane 3), was cleaved with *Eco*RI and fractionated by gel electrophoresis.

a) The ethidium bromide staining pattern.

b), c) After transfer to nitrocellulose, the genomic DNA was hybridized to radiolabelled DNA prepared from p(PP95)-P9.8 or p(PP95)-W17, respectively. Molecular size markers are indicated in kbp.

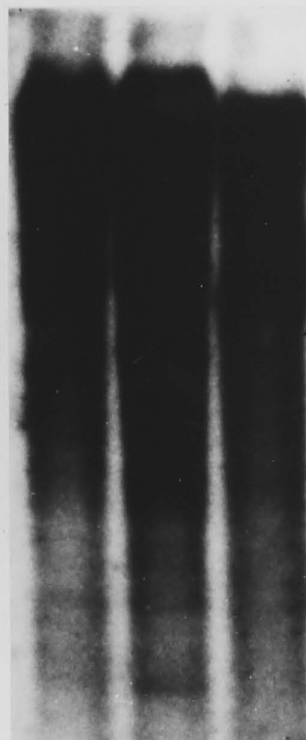
a)

1 2 3



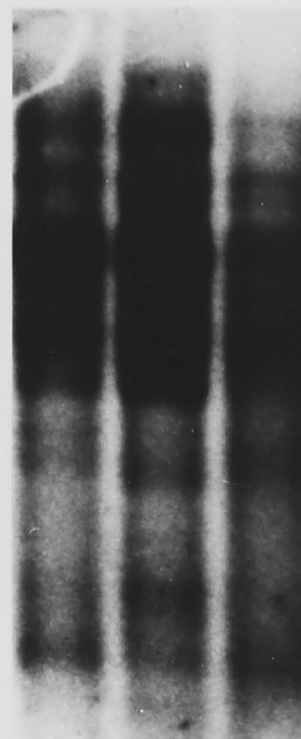
b)

1 2 3



c)

1 2 3



— 23.7

— 9.5

— 6.7

— 4.3

— 2.3

— 2.0

Figure 3.12: Localization of the cloned region to the *uncoordinated* complementation group.

a) Approximately 1  $\mu$ g genomic DNA extracted from Canton-S female adult heads (lane 1), and female pupae of B57/FM6 (lane 2) and B57/S54 (lane 3), was cleaved with *Eco*RI and separated by gel electrophoresis. After transfer to nitrocellulose, the DNA was hybridized to the fragments indicated in (b).\*

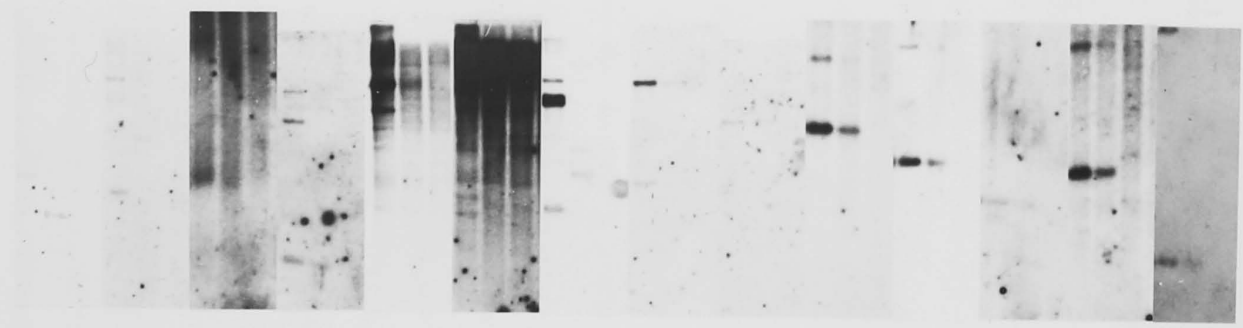
b) Genomic DNA divided into *Eco*RI fragments, illustrating the fragments used as hybridization probes. The cleavage site of the restriction enzyme *Pvu*II (p) is also indicated.

c) Coordinates in units of kbp.

\* In each case, the Southern blot was subsequently hybridised to DNA from the Adh gene in order to assess the relative amounts of genomic DNA tested.

noted  
heads  
was  
ar to  
\*  
used  
pVII  
NA

a)  
1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3





## CHAPTER FOUR

TRANSCRIPTIONAL ANALYSIS OF THE *UNCOORDINATED* COMPLEMENTATION GROUP

Studies described in the previous chapter resulted in the isolation and characterization of DNA sequences from the *uncoordinated* complementation group. In order to assess whether these sequences are of functional significance to the organism, or represent non-functional intergenic spacer DNA, the encoded transcription products were analysed and those required for production of the wildtype *uncoordinated* gene product identified. As the first clone isolated in the study was identified by hybridization to cDNA prepared from total adult head RNA (Chapter 3), it was anticipated that at least part of the cloned DNA would be transcribed.

The work described in this chapter also gives some insight into the mechanisms regulating expression of the *uncoordinated* gene. As initiation of transcription is one of the points of control of gene expression, a transcriptional analysis of the *uncoordinated* gene represents the initial step in understanding the regulation of expression of its gene product. In addition, this analysis allows the relationship between the unusual structural arrangement of DNA sequences at the *uncoordinated* complementation group and biological activity to be investigated.

#### 4.1 Identification of Transcripts Homologous to *uncoordinated* Sequences

In an attempt to identify transcripts encoded by the cloned DNA, a sensitive experimental approach was adopted, whereby the subcloned fragments were hybridized to Northern blots carrying RNA extracted from each major life stage of *Drosophila* (see Section 2.8.2). Adult head RNA was also tested since the isolation procedure of  $\lambda$  PP95 suggested that some part of its genomic

insert was producing a head-enriched transcript. Subcloned fragments from coordinates +23.6 to -11.0 were hybridized to both the total cellular and total cellular polyadenylated (poly A<sup>+</sup>) RNA populations. Reassociation experiments have demonstrated that only one-third of the polysomal RNA population is polyadenylated (Levy and Manning, 1981; Zimmermann *et al.*, 1980). Therefore, examination of the total cellular RNA population, which contains both the adenylated and non-adenylated RNA fractions, represents a first crude assessment of whether the *uncoordinated* complementation group produces non-adenylated transcripts.

Figure 4.1 shows that five rare transcripts were identified by *Eco*RI fragments C8, P1, P9.8, C9 and C20 (coordinates +12.8 to +9.4, 0 to +4, 0 to -2.75, -7.8 to -9.2 and -9.2 to -11.0). The transcripts, of sizes 4.35, 8.7, 9.0, 11.5 and 13.0kbp, were each detected in only the poly A<sup>+</sup> RNA fraction. It is possible that some of the larger transcripts represent partially processed precursors of the smaller species, as the total cellular polyadenylated fraction includes the nuclear RNA. However, this seems unlikely since each RNA species was produced in a temporally specific manner, and the larger transcripts were not homologous to those DNA sequences constituting the introns of the smaller transcripts.

*Eco*RI fragments C9 and P1 (coordinates -9.2 to -7.8 and 0 to +4) cross hybridize and are therefore similar in DNA sequences. The data from the Northern blot analysis is insufficient to determine if each fragment contains a coding region for the poly A<sup>+</sup> transcripts as it is possible that a non-coding fragment hybridizes to the transcript because of its homology with a fragment carrying a coding region. This also applies to fragments C20 and P9.8 (coordinates -9.2 to -11.0 and 0 to -2.75), which also cross-hybridize, albeit weakly.

A second class of RNA product was readily identified in the total cellular RNA fraction, but was much less abundant in the polyadenylated fraction (Figure 4.2). Only a small proportion of this RNA species fractionated into the poly A<sup>+</sup> class, possibly because either the transcript had a short polyadenylated 'tail', or was poly A<sup>-</sup>, but the internal sequences were rich in adenosine residues. The transcript is 7.3kbp in size and was identified by *Eco*RI fragments C5, C8 and C7 (coordinates +4.0 to +16.0). When hybridized to total cellular RNA, the C7 fragment also identified a number of smaller transcripts. However, these transcripts are not encoded by this region of the genome (see Section 4.5) and will not be considered further.

Although the number of RNA molecules per cell has not been estimated, it is apparent from the hybridization data that the polyadenylated transcripts are very rare. Figure 4.3(a) enables a direct comparison of the intensity of the hybridization signal produced by the transcripts encoded by DNA from the *uncoordinated* complementation group and those encoded by the alcohol dehydrogenase gene. In contrast, the 7.3kbp transcript is synthesized at sufficiently high levels to be identified in the total cellular RNA fraction of which only 3% is mRNA. Figure 4.3(b) compares the abundance of the 7.3kbp transcript with that of the *Adh* transcript.

The *Eco*RI fragments from coordinates -5.1 to -7.8 (W8 and W17) hybridized very intensely to several late larval poly A<sup>+</sup> transcripts.\* However, these transcripts have not been detected by other fragments in the region, and the W8 and W17 fragments contain insufficient DNA to encode the transcripts. It is possible that the transcripts originate from elsewhere in the genome and that each of these fragments is a member of a mobile element family, hybridizing to the transcriptional product of a related element located elsewhere. Transcripts synthesized by mobile elements are often polyadenylated and synthesized in a temporally specific manner (Schwartz *et*

\* Data not shown.

al., 1982), and fragments W8 and W17 have some features characteristic of mobile elements (see Sections 3.2 and 3.3). Alternatively, the other exons of the transcripts are located outside the cloned region.

Figure 4.4(b) summarizes the Northern blot hybridization data for sequences from the *uncoordinated* complementation group. The five poly A<sup>+</sup> transcripts each hybridized to the same genomic fragments, which span up to 28kbp. The coding sequences of the 7.3kbp species are confined to a region of approximately 12kbp which partially overlaps the coding region of the polyadenylated transcripts. This organization of coding regions suggests that a series of alternative processing events and/or DNA rearrangements gives rise to multiple transcripts.

#### 4.2 Temporal Specificity of Transcripts Homologous to *uncoordinated* Sequences

Each RNA species identified in the previous section is generated in a temporally specific manner (summarized in Figure 4.4(c)). For example, the 13.0kbp poly A<sup>+</sup> species was detected in early larvae, <sup>and in pupae</sup> whereas the 8.7 and 11.5kbp species were confined to the pupal stage. Although the 4.35 and 9.0kbp species have been identified in three life stages (early larvae, mid pupae and whole adults), quantitative differences were detected in the amount of each transcript present in each stage. The 9.0kbp species, present in similar amounts in early larvae and pupae, was barely detectable in whole adults, whereas the 4.35kbp transcript was present at very low levels in pupae and could not be detected by the fragments C8 and C20 (coordinates +9.4 to +12.8 and -9.2 to -11.0). The 4.35kbp transcript was the only poly A<sup>+</sup> species found in significant quantities in whole adults. The C5 fragment (coordinates +12.8 to +16.0) hybridized to the early larval poly A<sup>+</sup> transcripts but not to the RNA products from other life stages (Figures 4.2 and 4.4(b)). It is

unclear at present if the early larval transcripts are more readily identified, or if the C5 sequence contains coding regions specific for early larval products.

Production of the 7.3kbp transcript commenced early in larval life and continued throughout the remainder of the organism's development. It was most abundant during the adult stage in both the total cellular and total cellular poly A<sup>+</sup> RNA fractions.

#### 4.3 Spatial Distribution of Transcripts Homologous to *uncoordinated* Sequences

In contrast to the poly A<sup>+</sup> transcripts, the 7.3 kbp species was detected in the adult head RNA population (Figures 4.1 and 4.2). To ascertain if the 7.3kbp transcript is head specific, Northern blots carrying total cellular and total cellular poly A<sup>+</sup> RNA extracted from whole adults, adult heads and adult bodies were hybridized with the C7 fragment (coordinates +4 to +9.4). As illustrated in Figure 4.5, the 7.3kbp transcript is present in total cellular and total cellular poly A<sup>+</sup> RNA populations from both adult bodies and adult heads and is therefore not a head specific RNA product. Although the intensity of hybridization of the C7 fragment to total cellular RNA from adult bodies was less intense than to whole adults or adult heads, comparison of signal intensities after hybridization with *Adh* DNA demonstrates that less RNA from adult bodies was tested (Figure 4.5(b)).

#### 4.4 Direction of Transcription of RNA Products Homologous to *uncoordinated* Sequences

In order to determine the direction of transcription of each RNA product encoded by the *uncoordinated* complementation group, strand specific hybridization probes were prepared. Fragments P1, C7 and C8 (coordinates 0 to



+12.8) were each subcloned, in opposite orientations, into the *Eco*RI site of the vector pSP65 (Figure 4.6, see Table 3.1). Radiolabelled single stranded RNA products complementary to each strand of the DNA fragment were transcribed from the SP6 promotor using the bacteriophage SP6 RNA polymerase. These RNA probes were hybridized to Northern blots carrying both total cellular and/or total cellular poly A<sup>+</sup> RNA from the life stages described previously (Section 4.1).

RNA products complementary to DNA prepared from p(C7/65-1) and p(D3/65-1) hybridized to a transcript of 7.3kbp in the total cellular and total cellular poly A<sup>+</sup> fractions (Figure 4.7(a)), inferring that the 'anti-sense' RNA product is synthesized from this DNA strand, and the functional transcript from the opposite strand. This data, together with the Northern blot analysis (Figure 4.4), suggest that the 5' end of the 7.3kbp transcript lies between coordinates +12.8 to +16.0 (Figure 4.7(b)). The increased sensitivity of single-stranded hybridization probes allowed the identification of two previously undetected low abundance transcripts (Figure 4.7). One transcript, of 7.2kbp, was present only in embryos, whereas the other, of 6.6kbp was present in all life stages tested, but at lower abundance than the 7.3kbp transcript. Both of these RNA products were transcribed from the same DNA strand as the 7.3kbp species.

It has not proven possible to determine the direction of transcription of the poly A<sup>+</sup> transcripts described in section 4.1 using single stranded RNA hybridization probes prepared from either the P1 or C8 fragments. RNA products complementary to both strands of the P1 fragment hybridized to numerous transcripts (Figure 4.8), some of which were common to both strands of this, and the C8, fragment. However, these hybridization probes failed to detect the previously identified poly A<sup>+</sup> transcripts. There are several possible explanations for this. Firstly, full-length radio-labelled RNA



products may not have been synthesized, either because the genomic DNA contained a termination signal recognized by the SP6 RNA polymerase, or the coding region was distant from the polymerase promotor and the reaction inefficient. Secondly, these fragments may contain inverted repeated sequences. Formation of a stem and loop structure could prevent the hybridization of sequences in the stem and evidence for this type of structure comes from solution hybridization experiments using p(PP95)-P1 DNA. Under conditions unfavourable for the formation of DNA-DNA hybrids, an S1-nuclease resistant double-stranded DNA structure was formed (data not shown), suggesting that the fragment contains an inverted repeated element of approximately 500bp.

#### 4.5 Localization of Transcripts to the *uncoordinated* Complementation Group

Section 4.1 described the identification of transcripts homologous to DNA sequences localized to the *uncoordinated* complementation group. However, since the region contains repeated sequences (Section 3.2) it was necessary to establish that the transcriptional products are derived from *uncoordinated* and not from related sequences located elsewhere in the genome. This was achieved using genotypes heterozygous for overlapping deficiencies to produce individuals homozygous deficient for only the *uncoordinated* complementation group. The rationale was that RNA products transcribed from *uncoordinated* sequences would be absent from the RNA population isolated from organisms homozygous deficient for this region.

The remainder of the analysis of transcription of the *uncoordinated* complementation group is confined principally to the 4.35 and 7.3kbp transcripts identified in the whole adult RNA population. This life stage was chosen because of the simplicity of transcription relative to other stages and in constructing the stocks to be described, the appropriate genotypes could be readily distinguished as late imagos and adults.

#### 4.5.1 Construction of genotypes homozygous deficient for the *uncoordinated* complementation group

Genotypes homozygous deficient for the *uncoordinated* complementation group were constructed using the deficiency chromosomes B57 and S54. The breakpoints of each deficiency, as defined by contiguous lethal complementation groups (Schalet and Lefevre, 1976; Lefevre, unpublished) are: 19E1/19E2 and 19E8/F1 for B57 and 19E7/E8 and 20A1/A2 for S54 (Figure 4.9). Since the two deficiencies overlap at only the *uncoordinated* complementation group, offspring from mating individuals carrying these deficiency chromosomes were homozygous deficient for the DNA sequences from the *uncoordinated* complementation group. Genotypes homozygous deficient for *varied outspread* and *uncoordinated* were also constructed, using the B57 and Q539 chromosomes (Figure 4.9). Many of the progeny carrying homozygous deficiencies failed to eclose although they survived into the pupal stage and continued development past the wildtype eclosion time. Each rearranged chromosome was carried over a balancer chromosome marked by the dominant mutation *Bar*, whereas deficiency chromosomes were wildtype at this locus. Therefore, the mating protocol was established such that the homozygous deficiency genotypes could be selected as pupae on the basis of wildtype eyes (Figure 4.10). In addition, B57/FM6 and Canton-S female late pupae were collected for use as controls.

#### 4.5.2 Localization of adult transcripts

Total cellular and total cellular poly A<sup>+</sup> RNA were extracted from late pupae and newly eclosed adults that were homozygous deficient for *uncoordinated*, and also both *uncoordinated* and *varied outspread*. Northern blot hybridization analysis was used to test for the presence of the 4.35 and 7.3 kbp adult transcripts (Section 4.1) in these preparations.

In wildtype organisms, the C7 fragment (coordinates +4.0 to +9.4) hybridized to the 7.3kbp transcript and numerous smaller transcripts in the total cellular RNA population (Section 4.1). However, hybridization of this fragment to Northern blots carrying total cellular RNA from the deficiency genotypes B57/Q539 and B57/S54 failed to detect the 7.3kbp transcript (Figure 4.11(a)). In contrast, the smaller transcripts were present in the RNA population of both deficiency genotypes. The adult 4.35kbp poly A<sup>+</sup> transcript was detected by several wildtype fragments (Section 4.1), including P1 and C8 (coordinates 0 to +4 and +9.4 to +12.8). Hybridization of either of these fragments to Northern blots carrying total cellular polyadenylated RNA prepared from the deficiency genotypes failed to detect the 4.35kbp transcript. The C8 fragment also detects 7.3kbp transcript in wildtype organisms (Figure 4.2) and as illustrated in Figure 4.11, neither transcript was detected in the RNA populations from the deficiency genotypes. Hybridization of each filter to <sup>32</sup>P-labelled *Adh* DNA confirmed the presence of RNA in each lane.

These results demonstrate that the 7.3 and 4.35kbp adult transcripts detected by the fragments illustrated in Figures 4.1 and 4.2 are indeed transcribed from the DNA sequences absent in the B57/S54 genotype. The *uncoordinated* locus is the only vital complementation group in this deficiency, thus localizing the transcripts to this complementation group. In contrast, the presence of small transcripts homologous to the C7 fragment in the RNA population of the deficiency genotypes (Figure 4.11(a)) suggests that these species are transcribed from sequences located elsewhere in the genome.

#### 4.6 Conclusions

Six transcripts have been identified with homology to sequences from the *uncoordinated* complementation group. Five of these transcripts, of 4.35, 8.7, 9.0, 11.5 and 13.0kbp, are confined to the polyadenylated RNA fraction and their expression is temporally specific. In contrast, a 7.3kbp transcript, which may belong to the non-adenylated class of RNA is present throughout the organism's life, although the species detected in embryos is slightly smaller. This transcript is also produced in the adult head, although not exclusively. All six RNA products appear to be transcribed from the same genomic region, indicating that a complex series of alternative RNA processing events and/or DNA rearrangements is occurring at this gene. Transcripts of different sizes are generated that are temporally and spatially regulated, and compartmentalized into different classes of RNA.

The 4.35 and 7.3kbp adult transcripts were localized to the *uncoordinated* complementation group, suggesting that two transcripts code for at least part of the *uncoordinated* gene product. In addition, the 5' region of the 7.3kbp transcript appears to lie between coordinates +12.8 to +16.0.

Figure 4.1: Identification of polyadenylated transcripts homologous to *uncoordinated* sequences.

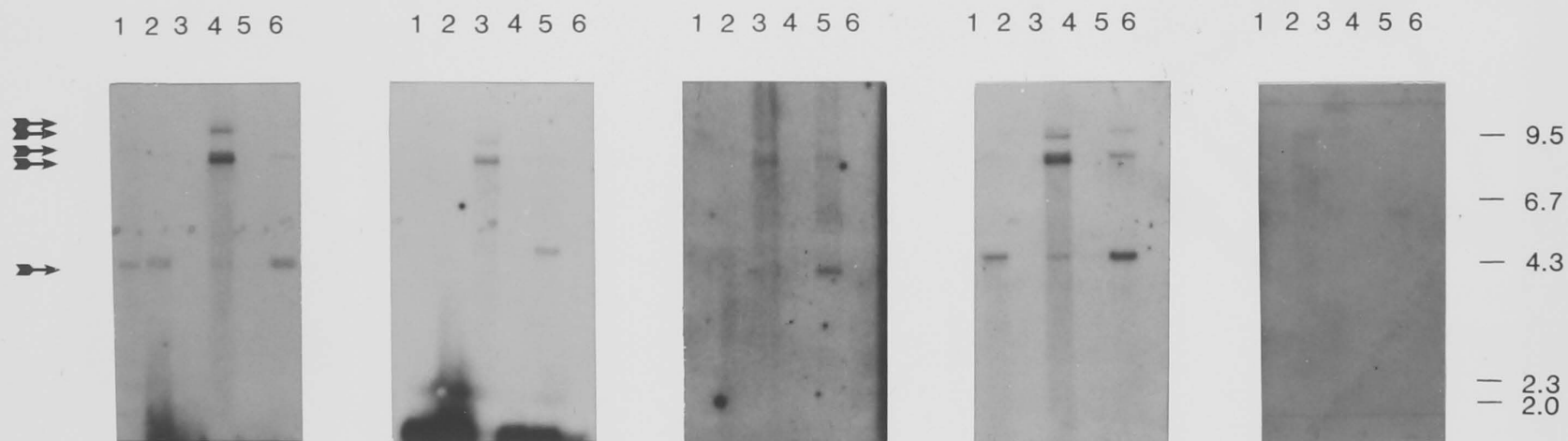
a) Hybridization of the fragments indicated in (b) to Northern blots carrying total cellular polyadenylated RNA: lane 1, whole adult; lane 2, adult head; lane 3, mid pupae; lane 4, late larvae; lane 5, early larvae; lane 6, mid embryos. Arrows show positions of transcripts. Molecular size markers are indicated in kbp.

b) Cloned DNA divided into *EcoRI* fragments, illustrating the fragments used as hybridization probes. The cleavage site of the restriction enzyme *PvuII* (p) is also shown.

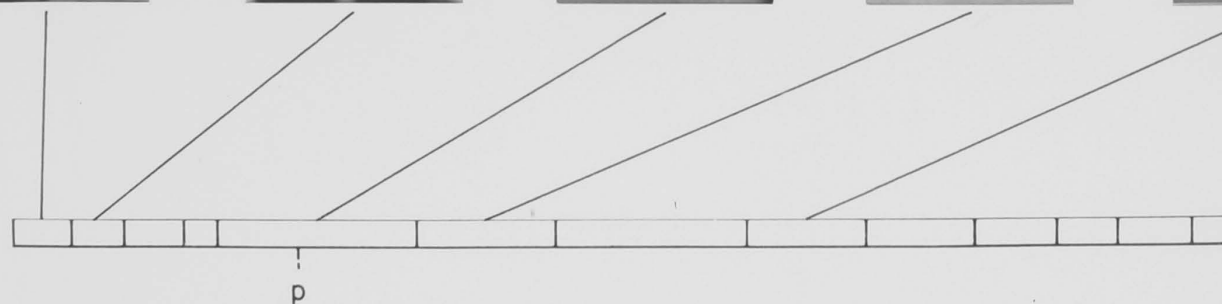
c) Coordinates in units of kbp.

is to  
rying  
head;  
mid,  
s are  
sed as  
I (p)

a)



b)



c)

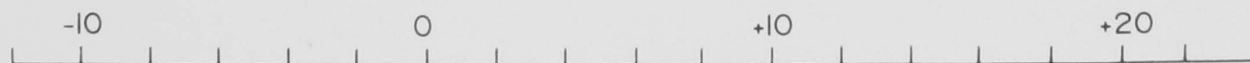




Figure 4.2: Identification of the 7.3 kbp transcript homologous to *uncoordinated* sequences.

a) Hybridization of the genomic fragments indicated in (b) to Northern blots carrying total cellular and total cellular polyadenylated RNA. Total cellular RNA: lane 1, whole adults; lane 2, adult heads; lane 3, mid pupae; lane 4, late larvae; lane 5, early larvae, lane 6, mid embryos. Total cellular polyadenylated RNA: lane 7, whole adults; lane 8, adult heads; lane 9, mid pupae; lane 10, late larvae; lane 11, early larvae; lane 12, mid embryos. The arrow shows the position of the 7.3 kbp transcript. Molecular size markers are indicated in kbp.

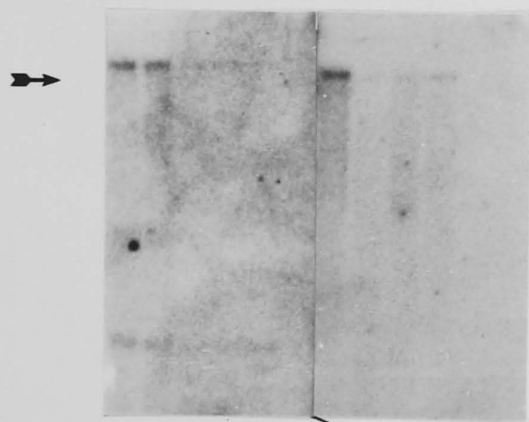
b) Cloned DNA divided into *EcoRI* fragments, illustrating the fragments used as hybridization probes. The cleavage site of the restriction enzyme *PvuII* (p) is also indicated.

c) Coordinates in units of kbp.

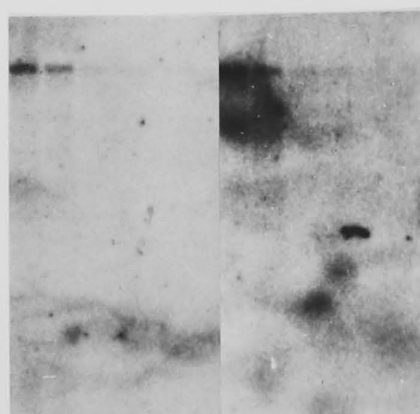
to  
lots  
lar  
4,  
lar  
mid  
The  
ers  
(p)  
as

a)

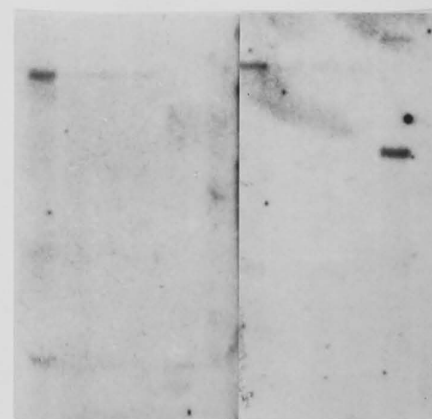
1 2 3 4 5 6 7 8 9 10 11 12



1 2 3 4 5 6 7 8 9 10 11 12

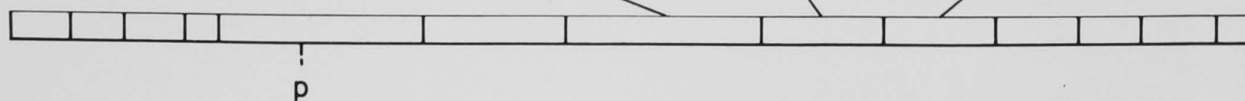


1 2 3 4 5 6 7 8 9 10 11 12



— 9.5  
— 6.7  
— 4.3  
  
— 2.3  
— 2.0

b)



c)

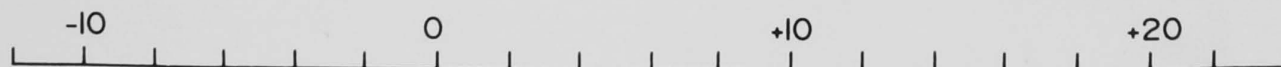


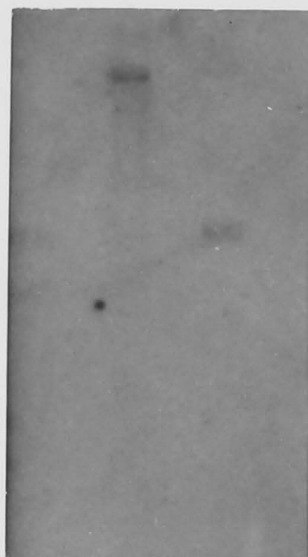
Figure 4.3: Comparison of the abundance of the 4.35 and 7.3kbp transcripts from the *uncoordinated* complementation group with that of the *Adh* transcripts.

a) Hybridization of the P1 fragment (left hand panel) and *Adh* DNA (right hand panel) to Northern blots carrying total cellular polyadenylated RNA.

b) Hybridization of the C7 fragment (left hand panel) and *Adh* DNA (right hand panel) to Northern blots carrying total cellular RNA. The RNA was extracted from: lane 1, whole adults; lane 2, adult heads; lane 3, mid pupae; lane 4, late larvae; lane 5, early larvae; lane 6, mid embryos. Molecular size markers are indicated in kbp.

a)

1 2 3 4 5 6



— 9.5 —

— 6.7 —

— 4.3 —

— 2.3 —

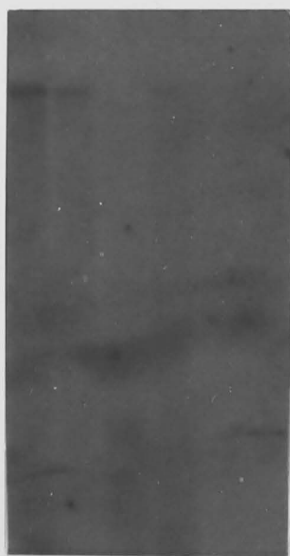
— 2.0 —

1 2 3 4 5 6



b)

1 2 3 4 5 6



— 9.5 —

— 6.7 —

— 4.3 —

— 2.3 —

— 2.0 —

1 2 3 4 5 6

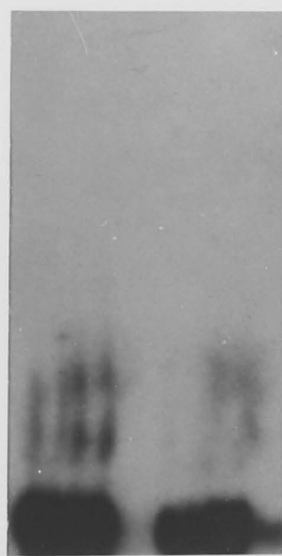


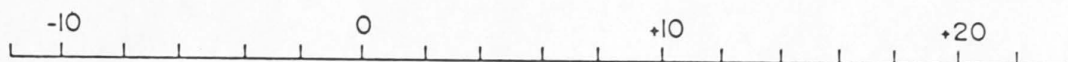
Figure 4.4: Summary of the organization and temporal specificity of the RNA species transcribed from the sequences of the *uncoordinated* complementation group.

a) Coordinates in units of kbp.

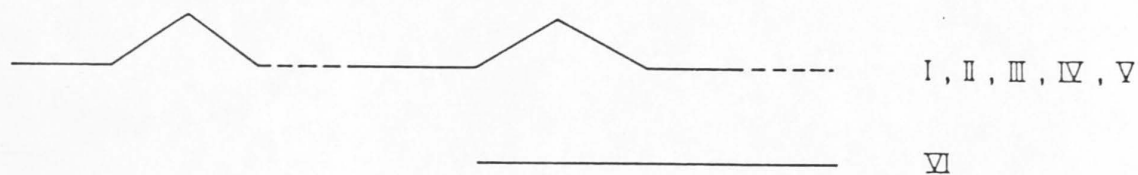
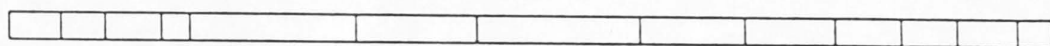
b) The genomic DNA divided into *EcoRI* fragments. The distribution of coding regions for each mRNA transcribed from the cloned sequences of the *uncoordinated* complementation group as determined by Northern blot analysis is indicated. The coding — , noncoding  $\Lambda$  , and possible further coding --- regions are shown.

c) Time of expression of each mRNA species transcribed from *uncoordinated* sequences. Each stage examined is listed on the abscissa. Thick and thin lines indicate a relatively strong or weak hybridization signal, respectively. I, II, III, IV, V and VI indicate transcripts of sizes 4.35, 8.7, 9.0, 11.5, 13.0 and 7.3 kbp, respectively.

a)



b)



c)

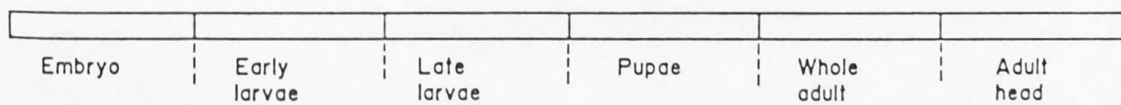
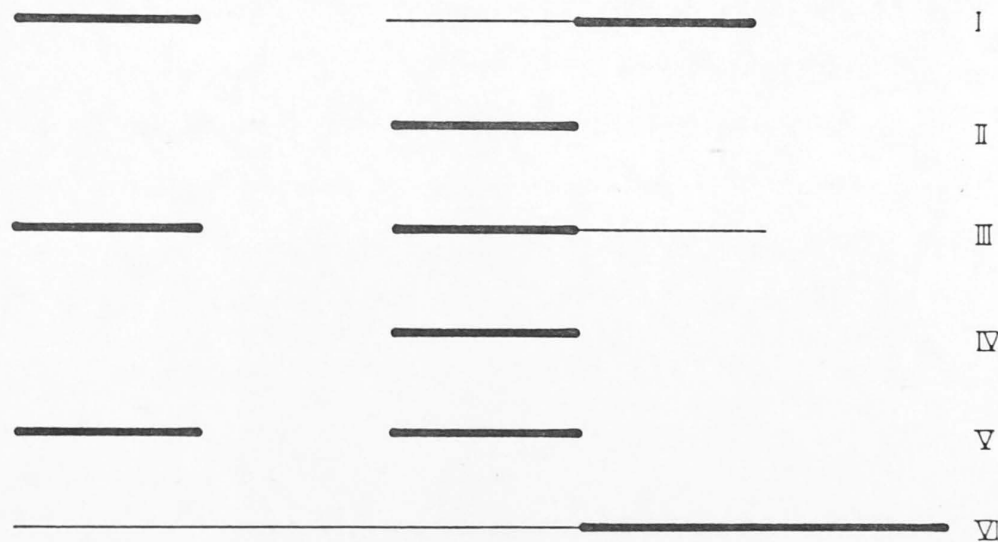




Figure 4.5: Spatial distribution of the 7.3 kbp transcript.

a) Hybridization of the C7 fragment to Northern blots carrying total cellular RNA extracted from: lane 1, whole adults; lane 2, adult heads; lane 3, adult bodies, and total cellular polyadenylated RNA extracted from: lane 4, whole adults; lane 5, adult heads; lane 6, adult bodies. The arrow indicates the location of the 7.3 kbp transcript.

b) Hybridization of DNA from the *Adh* gene to the same filter as in (a). Molecular size markers are indicated in kbp.

a)

1 2 3 4 5 6



— 9.5 —

— 6.7 —

— 4.3 —

— 2.3 —

— 2.0 —

b)

1 2 3 4 5 6

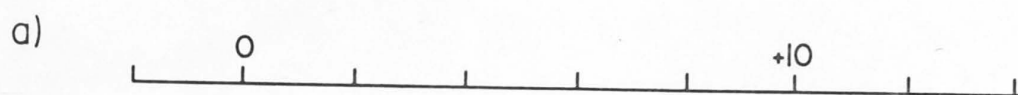


Figure 4.6: Direction of transcription of the single stranded RNA product complementary to each pSP65 recombinant plasmid.

a) Coordinates in units of kbp.

b) Genomic fragments subcloned into the pSP65 vector (see Table 3.1).

c) Cloned DNA divided into *Eco*RI fragments and indicating the 5'-3' orientation of the RNA product complementary to each subclone.



b)

PI C7 D3

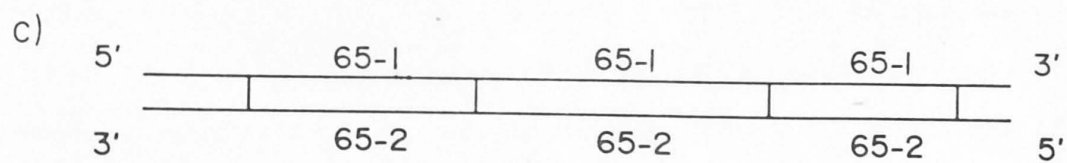


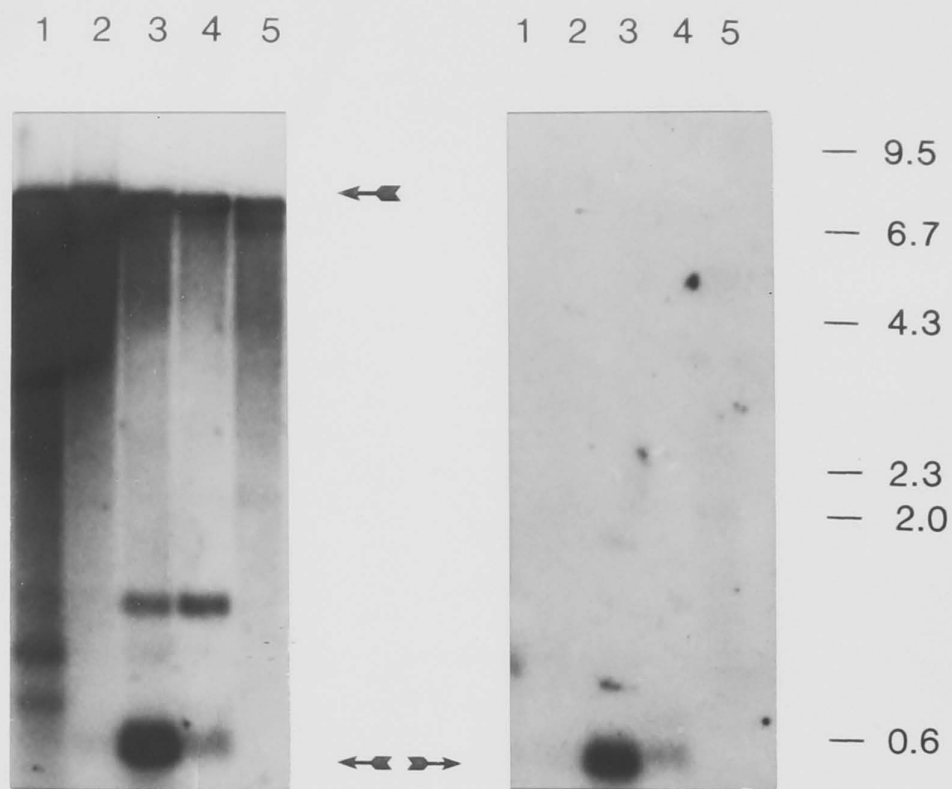
Figure 4.7: Direction of transcription of the 7.3 kbp transcript.

a) Hybridization of radiolabelled DNA from p(D3/65-1) (left panel) and p(D3/65-2) (right panel) to Northern blots carrying total cellular polyadenylated RNA extracted from: lane 1, whole adults; lane 2, mid pupae; lane 3, late larvae; lane 4, early larvae; lane 5, mid embryos. Arrows indicate the positions of 7.3 kbp transcript (top) and transcripts identified by both hybridization probes (bottom). Molecular size markers are indicated in kbp.

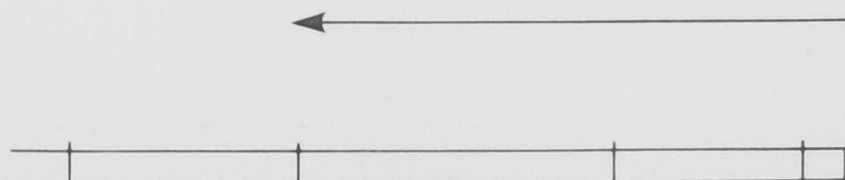
b) The direction of transcription of the 7.3 kbp transcript relative to the cloned DNA.

c) Coordinates in units of kbp.

a)



b)



c)

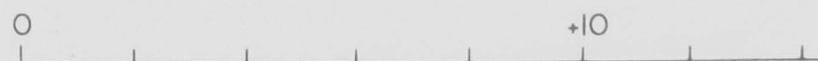




Figure 4.8: Direction of transcription of the polyadenylated transcripts.

Hybridization of radiolabelled DNA from p(P1/65-1) (a) and p(P1/65-2) (b) to Northern blots carrying total cellular polyadenylated RNA extracted from: lane 1, whole adults; lane 2, mid pupae; lane 3, late larvae; lane 4, early larvae; lane 5, mid embryos. Arrows indicate the positions of those RNA species identified by both hybridization probes. Molecular size markers are indicated in kbp.

a)

1 2 3 4 5

9.5 —

6.7 —

4.3 —

2.3 —

2.0 —

0.6 —



b)

1 2 3 4 5

← →

← →

← →

← →

← →



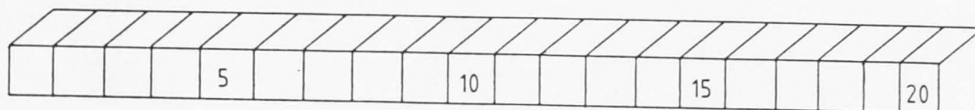
Figure 4.9: Breakpoints of chromosomal deficiencies relative to the polytene chromosome.

a) Schematic diagram of the polytene X chromosome showing the chromosomal divisions.

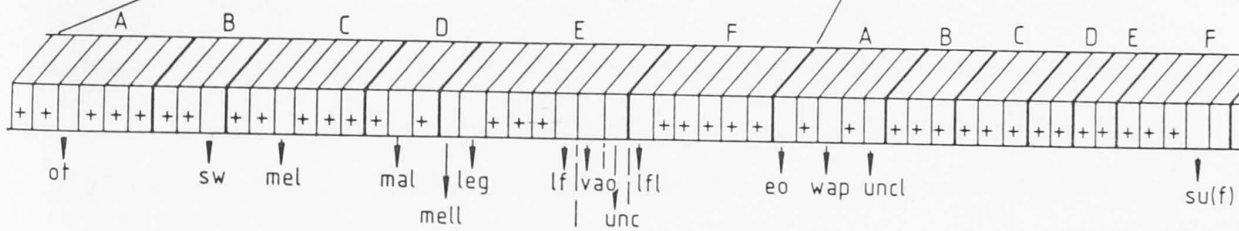
b) Divisions 19 and 20 illustrated in detail, showing the subdivisions (lettered) and each complementation group. The mapped visible mutations (see legend to Figure 1.3) and lethal (+) complementation groups are shown.

c) Breakpoints of each deficiency. Hatched regions indicate the chromosome.

a)



b)



c)

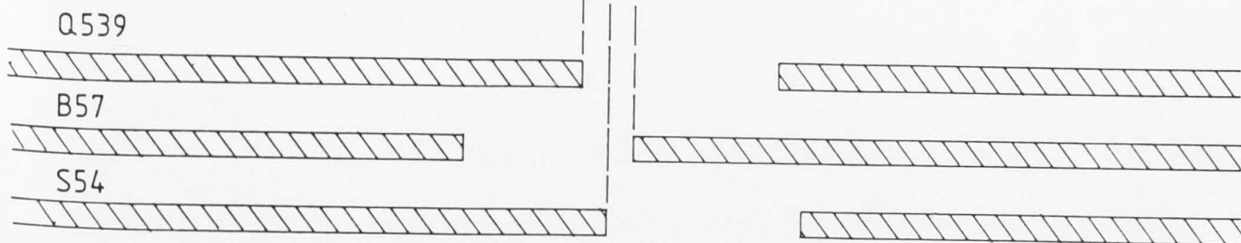


Figure 4.10: Construction of genotypes homozygous deficient for the *uncoordinated* complementation group. The parental females carried the S54 deficiency chromosome and the FM6 X chromosome balancer, carrying the dominant marker mutation, *Bar*. The eyes were phenotypically half-*Bar*. Males carried the B57 deficiency chromosome and the Y chromosome,  $y^{+}Ymal^{106}$ , and were wildtype with respect to eye phenotype. Individuals homozygous deficient for the *uncoordinated* complementation group were identified on the basis of their wildtype eyes in the pupal stage. At this stage the organisms could be sexed and the eye phenotype observed.

$G_0$ :                      B57/ $y^+Ymal^{106}$                        $\times$                       S54/FM6

eye phenotype:                      Wildtype                                      Bar



$G_1$ :                      S54/ $y^+Ymal^{106}$  : FM6/ $y^+Ymal^{106}$  : B57/S54 : B57/FM6

eye phenotype:                      Wildtype : Bar : Wildtype : Bar



Collected at  
late pupae



Figure 4.11: Localization of adult transcripts to the *uncoordinated* complementation group.

a) Hybridization of the C7 fragment to Northern blots carrying total cellular RNA extracted from: lane 1, Canton-S adults, and whole pupae of lane 2, B57/FM6, lane 3, B57/Q539, lane 4, B57/S54, lane 5, *unc*<sup>16.3.212</sup>/*unc*<sup>16.3.212</sup>.

b) Hybridization of the C8 fragment to Northern blots carrying total cellular polyadenylated RNA extracted from the same genotypes as in (a). The arrows indicate the positions of the 7.3 (a) and (b) and 4.35 (a) kbp transcripts. Molecular size markers are indicated in kbp.

dimethyl  
cellular  
lane 2,  
212.  
cellular  
arrows  
cripts.

a)

1 2 3 4 5

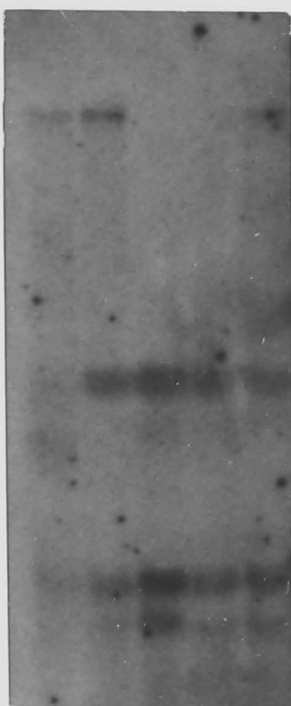
9.5 —

6.7 —

4.3 —

2.3 —

2.0 —



b)

1 2 3 4 5



## CHAPTER FIVE

ANALYSIS OF MUTATIONS IN THE UNCOORDINATED GENE

One strategy for studying the organization, regulation and function of a gene is to analyse its mutated alleles. This approach has been exploited genetically to predict the spatial and temporal specificity of gene expression and to define interactions between genes. In some instances, important insights have been gained about role of the gene in development. For example, the homeotic genes of *Drosophila* have been shown to control the identity of body segments (Ouweneel, 1976). At other loci, such as *white*, structural and regulatory domains have been defined (reviewed in Judd, 1976).

This approach has also been successfully employed in the molecular analysis of genes. It has been possible to correlate molecular and genetic maps, identify the location of the mutations at the DNA sequence level and examine the molecular nature of mutational lesions, (for example Bender *et al.*, 1983(a); Zachar and Bingham, 1982). In several cases, genetic predictions have been tested at the molecular level. For example, the genetic data predicted that the *Ultrabithorax* gene product was active in the third thoracic segment (Lewis 1978), and this has been confirmed by *in situ* localization of the *Ultrabithorax* transcripts in early embryos (Akam, 1983).

Several mutant alleles of the *uncoordinated* locus have been induced by radiation and chemical mutagenesis (Table 5.1). In this study, the DNA sequence organization and transcriptional products of seven mutant *uncoordinated* alleles were analysed. It was anticipated that such an analysis would confirm that the *uncoordinated* gene had been cloned and provide an estimate of its size.

### 5.1 Analysis of DNA Sequences from *uncoordinated* Mutant Alleles

Seven mutants, phenotypically unable to coordinate leg movement, have been mapped to the *uncoordinated* complementation group (Lefevre, unpublished). Each mutation is semi-lethal in the homozygous state and mutants survive into the late pupal stage with only a small percentage eclosing. Because DNA from the *uncoordinated* complementation group contains repeated sequences, and there are strain differences in restriction enzyme cleavage sites, the DNA from homozygous mutant alleles was examined to avoid ambiguities resulting from the presence of a wildtype chromosome. Homozygous mutant alleles were constructed using the mating protocol outlined in Figure 5.1. Matings were established to allow the homozygous mutant genotypes to be identified on the basis of eye characteristics in the pupal stage. The homozygous mutants were collected as males for two alleles (Figure 5.1(b)) and females for the remaining alleles (Figure 5.1(a)).

#### 5.1.1 Southern blot hybridization analysis of *uncoordinated* mutant alleles

Genomic DNA was extracted from the two wildtype strains, Canton-S and M56i, and each *uncoordinated* mutant homozygote. Southern blots carrying *Eco*RI and *Hind*III cleaved DNA from each genotype were hybridized with subcloned fragments from coordinates +23.5 to -11.0 (Figure 3.1). Restriction enzyme cleavage maps were inferred from the hybridization data.

The restriction maps of the four mutant *uncoordinated* alleles derived from the M56i chromosome, GE230, DC803, S95 and S86, were identical (Figure 5.2) but different from that of M56i. Hybridization of the 4.0kbp P1 fragment (coordinates 0 to +4.0) to *Eco*RI cleaved M56i DNA produced a strong hybridization signal to a fragment of 4.0kbp, and a much weaker signal to fragments of 4.2 and 5.2kbp. However, when the P1 fragment was hybridized with *Eco*RI digested DNA from the *uncoordinated* mutants, the hybridization

signal of fragments 4.2 and 5.2kbp was intense and that of a third fragment of 3.8kbp was barely detectable (Figure 5.2(a)). As shown previously in Figure 3.1, 1.37kbp C9 fragment (coordinates -7.8 to -9.2) hybridizes with P1. When this sequence was tested against M56i DNA, the hybridization signals to fragments of 4.0kbp, 4.2kbp and 5.2kbp were more equal in intensity.\* However, hybridization of this fragment to the DNA of each *uncoordinated* mutant revealed a signal of about equal intensity to three genomic bands of sizes 5.2, 4.2 and 3.8kbp.

The data suggest that the same DNA rearrangement has occurred in each mutant. Sequences related to the C9 fragment are relocated to sites normally containing sequences related to the P1 fragment and conversely, the P1 type sequences to sites that normally contain C9 sequences (Figure 5.2(b)). In addition, a small deletion of approximately 200bp has occurred within the 4.0kbp *EcoRI* fragment of M56i to produce a 3.8kbp fragment in the mutant DNA. Figure 5.2(c) shows one arrangement of genomic fragments that is consistent with the hybridization data. This arrangement would have resulted from an inversion with breakpoints within the 1.37kbp and 4.2kbp fragments (coordinates -7.8 to -9.2 and 0 to +4.2, respectively), and in generating this chromosomal rearrangement, a small deletion may have occurred.

It is difficult to assess if the proposed changes in DNA sequence organization of the four mutant alleles are responsible for the *uncoordinated* mutant phenotype. This is partly due to the unexpected observation that all four mutant alleles appear to contain very similar DNA rearrangements despite having been induced with different mutagens. A direct comparison of the mutant and parental DNA reduces the possibility of confusing strain differences with changes due to *uncoordinated* mutations. However, it remains a formal possibility that the difference between the mutant and parental stocks in the intensity of the hybridization signal of the two fragments

\* Data not shown.

above, is not causally related to the *uncoordinated* mutant phenotype. For example, a restriction fragment polymorphism in the population may have been selected during the isolation of *uncoordinated* mutants. If this were the case, the small deletion would appear to be responsible for the *uncoordinated* mutant phenotype.

Southern blot hybridization of DNA from the *uncoordinated* mutant W5 (Figure 5.3(a)) revealed an insertion of 300-400bp in the B46 fragment (coordinates +17.4 to +20.3). Hybridization of fragment B46 to *Hind*III cleaved W5 DNA further suggests that the insertion has occurred in the 1.35kbp *Eco*RI-*Hind*III fragment (coordinates +16.0 to +17.4). A deletion has also occurred in the region from coordinates -7.8 to -8.7, <sup>the *Canten-S*</sup> <sup>i.e. the C9/C20 fragments</sup> resulting in a single 2.55kbp *Eco*RI genomic fragment instead of two of sizes 1.37 and 1.8kbp (Figure 5.3(a)). The position of these two lesions on the restriction map is summarized in Figure 5.3(b). The *Drosophila* strain carrying the *w* chromosome, from which the W5 mutant was generated (Table 5.1), was not available for analysis and comparison.

Hybridization of the C20 fragment <sup>*Canten-S*</sup> (coordinates -9.2 to -11.0) to *Eco*RI cleaved DNA of the 16.3.212 mutant suggests that an insertion has occurred within this fragment (Figure 5.4(a)). Hybridization of this, and the adjacent C9 fragment <sup>*Canten-S*</sup> (coordinates -7.8 to -9.2) to *Hind*III cleaved 16.3.212 DNA localized the insertion to within 0.5kbp of the *Hind*III site and demonstrated that a small deletion had occurred in the C9 fragment (Figure 5.4(b)). The 16.3.212 mutant was derived from the Oregon-K wildtype strain (Table 5.1), which was not available. Therefore the DNA from the 16.3.212 allele could not be compared with that of the wildtype allele.

Hybridization analysis of DNA from the 27E2 *uncoordinated* mutant suggests that P1 fragment (coordinates 0 to +4.0) has been amplified (Figure 5.5(a)). It is not possible to determine if several fragments of different lengths are



amplified because of the strong hybridization signal in the 4.0kbp region, although regions adjacent to fragment P1 do not appear to be amplified (data not shown). It is not known from which *Drosophila* strain the 27E2 mutant was generated, and therefore the parental strain was unavailable for analysis.

Quantitative increases in gene copy number have been described for several gene systems, which differ in the mechanisms by which the increase is generated. For example, rRNA genes in amphibian oocytes are amplified in 'head to tail' tandem arrays (Brown and Dawid, 1968) whereas dihydrofolate reductase genes are amplified as extra-chromosomal elements conferring methotrexate resistance on cultured mammalian cells (reviewed in Schimke, 1984). The chorion genes are the best characterized example of gene amplification in *Drosophila* (Spradling and Mahowald, 1980). Additional copies of these genes are synthesized by the formation of multiple replication forks (Spradling, 1981). The available data do not allow determination of the organization of the amplified sequence in the 27E2 allele.

While the DNA sequences of the mutant alleles S95, S86, DC803 and GE230 were compared directly to that of the parental strain, M56i, the parental strains of the 16.3.212, W5 and 27E2 alleles were not available for analysis. Although for each of these alleles, alterations in the organization of the *uncoordinated* restriction fragments were correlated with the mutant phenotype, it remains a formal possibility that the described alterations result from strain differences. It is important to note that Southern blot hybridizations revealed no difference between DNA sequences from the *uncoordinated* complementation group in Canton-S adult heads and pupae (Figure 3.12). It is therefore considered unlikely that the changes detected in the genomic DNA prepared from mutants in the pupal stage were due to differential representation of the sequences in polytene chromosomes. It is probable, therefore, that the observed sequence alterations are responsible for the *uncoordinated* mutant phenotype.

The lesions of each of the seven *uncoordinated* mutant alleles are at least partly associated with sequences homologous to the 4.0kbp P1 fragment (summarized in Figure 5.8). In some cases, these sequences are the only detectable sites of change in DNA sequence organization, indicating that these sequences are highly susceptible to mutagenesis. Subdivision 19E8 is highly susceptible to irradiation-induced chromosome breakage (for example, Lefevre, 1981; Figure 1.2). Analysis of the sequences in which the breakages occurred would enable an assessment of any correlation between the DNA sequences containing the breakpoints of the chromosomal rearrangements and the location of P1-related sequences.

#### 5.1.2 Cytology of *uncoordinated* mutants

In order to assess if any of the *uncoordinated* mutations resulted from gross chromosomal rearrangements, each mutant chromosome was subjected to cytological examination. Because of the unexpected observation that the four M56i derived mutants had identical restriction maps at the *uncoordinated* complementation group, it was difficult to assess if the proposed changes in their restriction maps were responsible for the *uncoordinated* phenotype. In addition, cytological examination of these alleles by Lefevre (unpublished) indicated that the S95 mutation was the result of a small chromosome rearrangement which was not detected in GE230, DC808 or S86 mutants. It was therefore important to test the mutants for cytologically detectable chromosomal rearrangements.

Polytene chromosomes were prepared from salivary glands of third instar larvae heterozygous for the *uncoordinated* mutant allele and a wildtype Canton-S X chromosome. Chromosomal rearrangements would disrupt pairing of the two chromosomes and would be visualised as discontinuities to the uniform appearance of the chromosome. The type of structure formed would be dependent on the nature of the rearrangement.

Each *uncoordinated* mutant allele was examined (Figure 5.6). Structural abnormalities were not apparent in any of the *uncoordinated* mutants tested, indicating that if rearrangements had occurred, they were too small to be detected cytologically.

## 5.2 Transcriptional Analysis of *Uncoordinated* Mutant Alleles

The transcription products of each *uncoordinated* mutant allele were analysed to determine if the mutant lesions resulted in any alteration in the production or processing of RNA products. Because mutant homozygotes were collected as a mixture of late pupae and adults, the two previously identified adults transcripts of 4.35 and 7.3kbp (Figure 4.4), could be assayed.

The total cellular RNA population of each mutant was tested for the presence of the 7.3kbp transcript by Northern blot hybridization with the C7 *Eco*R1 fragment (coordinates +4 to +9.4). Figure 5.7(b) illustrates that full-length copies of this transcript are present in each mutant. Comparison of the intensities of the hybridization signals observed with mutant and Canton-S RNA demonstrates that approximately wildtype levels of the transcript were produced in each mutant. The filter was subsequently hybridized with *Adh* DNA to demonstrate that similar amounts of RNA from each genotype were tested.

Hybridization of the P1 fragment (coordinates 0 to +4.0) to Northern blots carrying the total cellular polyadenylated RNA population of each mutant failed to detect the 4.35kbp transcript (Figure 5.7(a)). However, four previously unidentified transcripts, of 7.0, 8.1, 11.5 and approximately 15.5kbp, were found in the W5 RNA population, suggesting that a disruption had occurred in the production and/or processing of the 4.35kbp transcript. The W5 poly A<sup>+</sup> transcripts have two distinctive properties. Firstly, the size of each transcript was a novel, discrete class that had not been previously detected. The new class of transcript may represent incorrectly, or partially

spliced, precursor products. Secondly, the W5 poly A<sup>+</sup> transcripts are produced in high abundance compared to the wildtype 4.35kbp transcript, with the 7.0kbp species present in greatest amounts. Comparison of the intensity of the hybridization signal from the P1 fragment with that from *Adh* DNA, demonstrated that approximately equal amounts of RNA from each allele was tested (Figure 5.7(a)). This suggested that one of the alterations to the DNA from the *uncoordinated* complementation group in W5 disrupted a region regulating the synthesis of the 4.35kbp transcript.

Figure 5.8 compares those fragments carrying coding sequences, as determined by Northern blot analysis, with those in which the DNA sequence organization has been altered in each *uncoordinated* mutant allele. DNA alterations are correlated with the location of the coding regions of the polyadenylated transcripts, but not with that of the 7.3kbp species. This suggests that the sequence changes are probably responsible for the absence of the 4.35kbp transcript. This comparison also allows an estimate of the limits of the *uncoordinated* gene within the cloned DNA. The *uncoordinated* gene appears to extend from coordinates +17.5 to at least -11.0. However, it is unknown whether sequences homologous to those within these coordinates, but located downstream of coordinate -11.0, are also part of this gene. It therefore encompasses at least 28kbp.

### 5.3 Conclusions

Alterations were found in the genomic DNA of seven *uncoordinated* mutants hybridizing to the cloned region of the *uncoordinated* complementation group. The correlation between DNA sequence alterations and the *uncoordinated* mutant phenotype provided additional evidence that the region under study represents at least part of the *uncoordinated* gene.

Production of the 4.35kbp poly A<sup>+</sup> transcript is disrupted in each of the mutants, although synthesis of the 7.3kbp species appears normal. Sequences homologous to the P1 fragment appear to be susceptible to mutagenesis, and the sequence alterations in these fragments are probably responsible for disrupting the production of the polyadenylated transcript. Regions have also been identified which may control the abundance, and possibly processing, of the polyadenylated transcripts. The absence of the 4.35kbp transcript from RNA populations of *uncoordinated* mutants also supports the proposal (Section 4.2.2) that this RNA species encodes at least part of the *uncoordinated* gene product. However, the role of the 7.3kbp species in the functioning of the *uncoordinated* gene remains unknown and this point is addressed in Chapter 6.

Analysis of *uncoordinated* mutants at the DNA and RNA levels confirms that the *uncoordinated* gene has been cloned and suggests that the gene extends over at least 28kbp, from coordinates +17.5 to at least -11.0.

Table 5.1

Origin of *uncoordinated* mutant alleles

| <u>Allele</u>         | <u>Parental Chromosome</u> | <u>Mutagen</u>       |
|-----------------------|----------------------------|----------------------|
| S86 <sup>*</sup>      | M56i                       | X-ray                |
| S95 <sup>*</sup>      | M56i                       | X-ray                |
| GE230 <sup>*</sup>    | M56i                       | X-ray                |
| DC803 <sup>*</sup>    | M56i                       | EMS                  |
| W5 <sup>∇</sup>       | w chromosome               | EMS                  |
| 16.3.212 <sup>#</sup> | Oregon-K                   | Neutron<br>radiation |
| 27E2 <sup>†</sup>     | unknown                    | unknown              |

<sup>\*</sup>Lefevre (unpublished)

<sup>∇</sup>Lifschytz and Falk, 1968

<sup>#</sup>Schalet and Lefevre, 1976 (kindly supplied by E. Munoz)

<sup>†</sup>Schalet (unpublished)



Figure 5.1(a): Construction of homozygous *uncoordinated* mutants, S95, S86, 27E2, 16.3.212 and W5.

The parental females carried the *unc* allele and a balancer chromosome carrying the dominant marker mutation, *Bar*. The eyes were phenotypically half-*Bar*. Males carried the *unc* allele and the Y chromosome,  $y^{+Ymal^{106}}$ , and were wildtype with respect to eye phenotype. The offspring of the mating were examined as late pupae. Females homozygous mutant at the *uncoordinated* locus were identified on the basis of wildtype eyes.

a)

G<sub>0</sub>:                      *unc/y<sup>+</sup>Ymal<sup>106</sup>*                      ×                      *unc/balancer*

eye phenotype:                      Wildtype                      Bar



G<sub>1</sub>:                      *unc/y<sup>+</sup>Ymal<sup>106</sup>* : *balancer/y<sup>+</sup>Ymal<sup>106</sup>* : *unc/unc* : *unc/balancer*

eye phenotype                      Wildtype                      :                      Bar                      : Wildtype:                      Bar



Collected at  
late pupae

Figure 5.1(b): Construction of homozygous *uncoordinated* mutants, GE230 and DC803.

The parental female was as described in the legend to Figure 5.1(a) and the parental male carried the Canton-S X and Y chromosomes. The offspring were examined as late pupae. Males homozygous mutant at the *uncoordinated* gene were identified on the basis of wildtype eyes.

b)

$G_0$ : CS/Y  $\times$  *unc*/balancer

eye phenotype: Wildtype Bar



$G_1$ : *unc*/Y : balancer/Y : *unc*/CS : CS/balancer

eye phenotype: Wildtype : Bar : Wildtype : Bar



Collected at  
late pupae

Figure 5.2: Restriction enzyme cleavage map of the *uncoordinated* mutant alleles, S95, S85, GE230 and DC803.

a) Approximately 1 µg genomic DNA extracted from whole pupae of each genotype was cleaved with *Eco*RI and fractionated by gel electrophoresis. After transfer to nitrocellulose, the DNA was hybridized to radiolabelled p(PP95)-P1 DNA. The following genotypes were tested: lane 1, M56i; lane 2, S95/S95; lane 3, S86/S86; lane 4, DC803/Y; lane 5, GE230/Y. Molecular size markers are indicated in kbp. \*

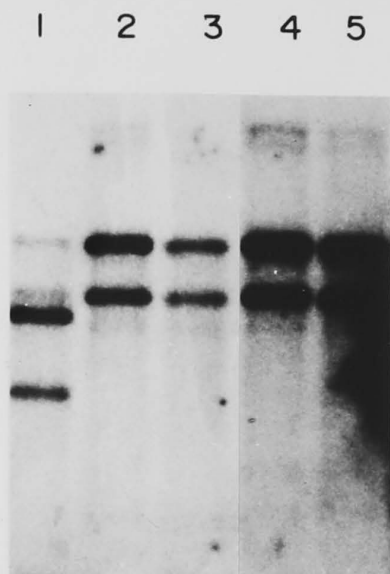
b) The proposed restriction enzyme cleavage map of the M56i parental strain. The cloned DNA is divided into *Eco*RI fragments and the location and nature of lesions observed in the derived mutant alleles are shown: ▲ the approximate breakpoints of the inversion; Δ a deletion of approximately 200bp; ▽ a duplicated fragment. Regions which hybridize to Canton-S fragments are indicated by the fragment name (see Figure 3.1). The cleavage sites of the restriction enzymes *Hind*III (H) and *Pvu*II (P) are indicated.

c) Proposed restriction enzyme cleavage map of the alleles S95, S86, GE230 and DC803. The DNA is divided into *Eco*RI fragments and regions with homology to the Canton-S fragments are indicated.

d) Coordinates in units of kbp.

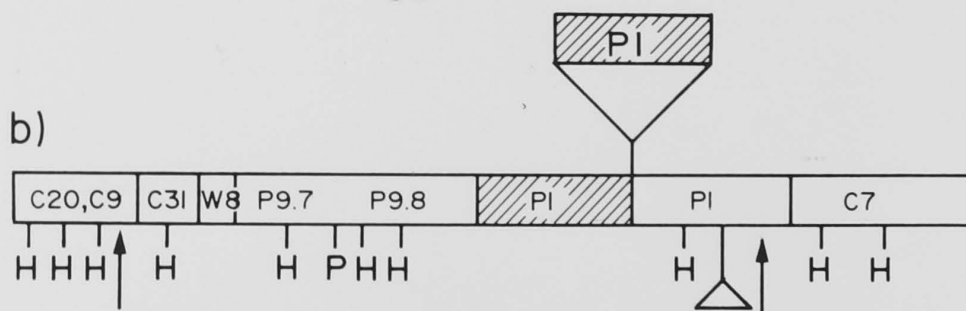
\* The hybridization signal at approximately 3.0 kbp in lane 1 is due to the presence of DNA with homology to the bacterial plasmids used as cloning vectors.

a)

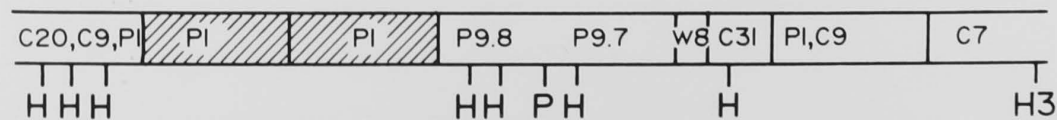


— 9.5  
— 6.7  
— 4.3  
  
— 2.3  
— 2.0

b)



c)





d)





Figure 5.3: Restriction enzyme cleavage map of the *uncoordinated* mutant allele, W5.

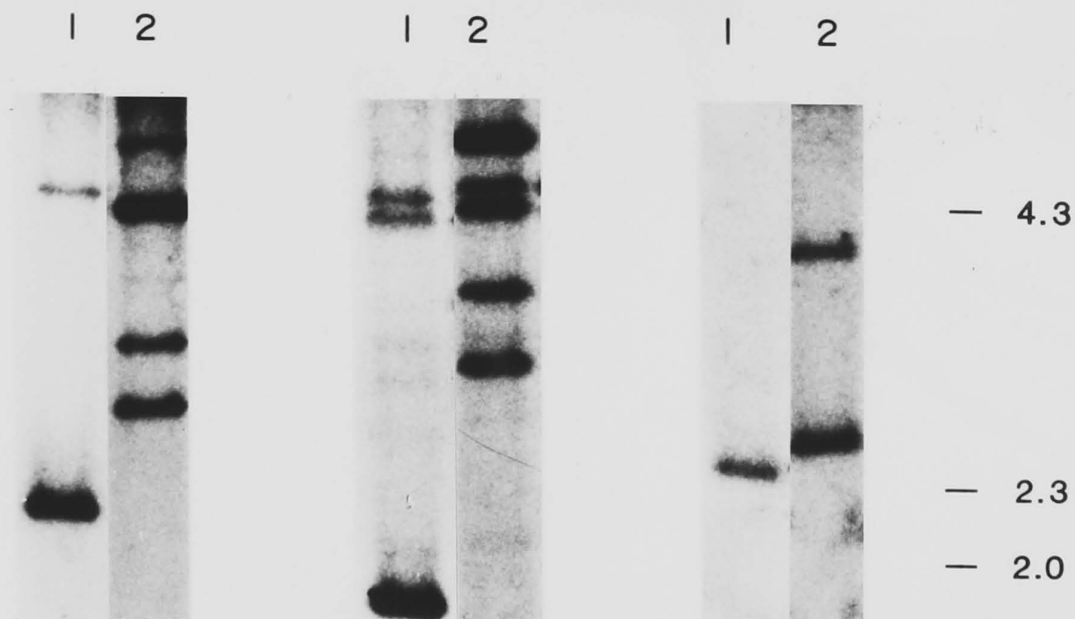
a) Approximately 1 µg genomic DNA, extracted from: lane 1, Canton-S female adult heads, lane 2, W5/W5 whole pupae, were cleaved with *Eco*RI and fractionated by gel electrophoresis. After transfer to nitrocellulose, the DNA was hybridized to radiolabelled DNA from p(MH26)-C20 (left panel), p(MH26)-C9 (middle panel) or p(RJ43)-B46 (right panel). Molecular size markers are indicated in kbp. \*

b) The proposed restriction enzyme cleavage map of the W5 allele. The DNA is divided into *Eco*RI fragments and the nature and location of the lesions observed in the W5 mutant are shown:  an insertion of 300-400 bp;  a small deletion between the *Hind*III and *Eco*RI sites. The cleavage sites of the restriction enzymes *Hind*III (H) and *Pvu*II (p) are indicated.

c) Coordinates in units of kbp.

\* The hybridization signal at approximately 3.0 kbp in lane 2 is due to the presence of DNA with homology to the bacterial plasmids used as cloning vectors.

a)



b)

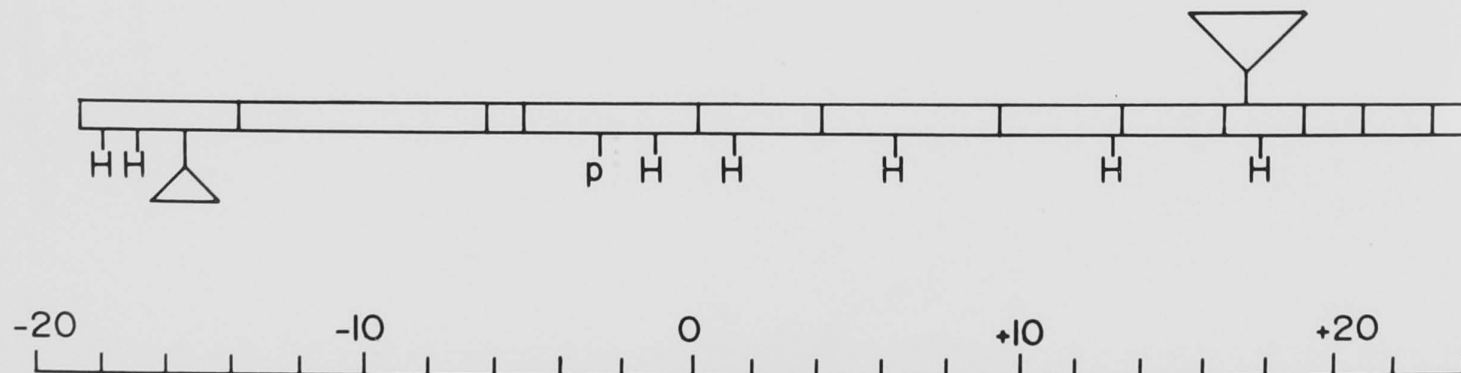




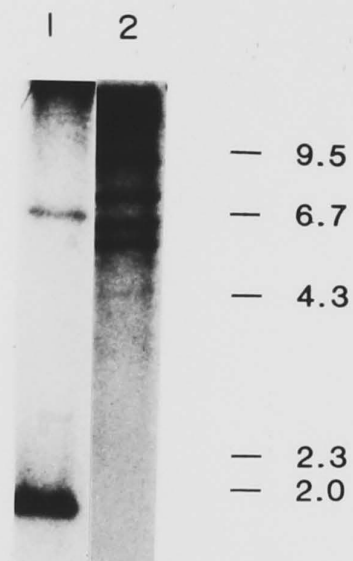
Figure 5.4: Restriction enzyme cleavage map of the *uncoordinated* mutant allele, 16.3.212.

a) Approximately 1 µg genomic DNA, extracted from: lane 1, Canton-S female adult heads, lane 2, 16.3.212/16.3.212 whole pupae, were cleaved with *Eco*RI and fractionated by gel electrophoresis. After transfer to nitrocellulose, the DNA was hybridized to radiolabelled DNA from p(MH26)-C20. Molecular size markers are indicated in kbp.

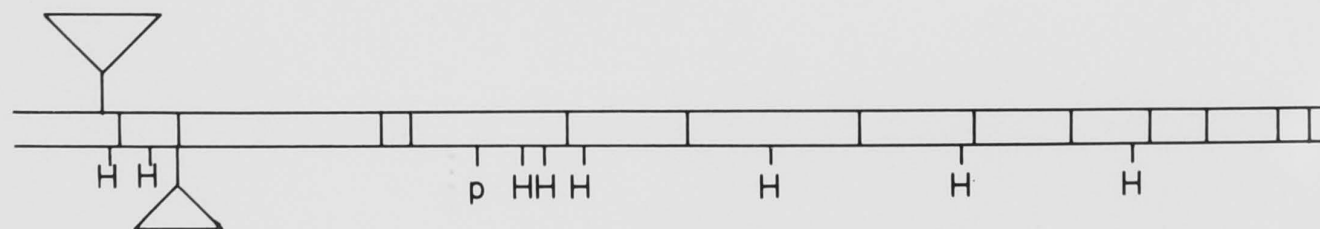
b) The proposed restriction enzyme cleavage map of the 16.3.212 allele. The DNA is divided into *Eco*RI fragments and the nature and location of the lesions observed in the mutant are shown:  an insertion;  a small deletion between the *Hind*III and *Eco*RI sites. The cleavage sites of the restriction enzymes *Hind*III (H) and *Pvu*II (p) are indicated.

c) Coordinates in units of kbp.

a)



b)



c)



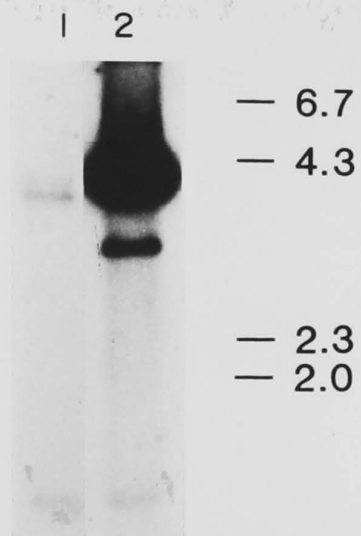
Figure 5.5: Restriction enzyme cleavage map of the *uncoordinated* mutant allele, 27E2.

a) Approximately 1 µg genomic DNA, extracted from: lane 1, Canton-S female adult heads, lane 2, 27E2/27E2 whole pupae, were cleaved with *Eco*RI and fractionated by gel electrophoresis. After transfer to nitrocellulose, the DNA was hybridized to radiolabelled DNA from p(PP95)-P1. Molecular size markers are indicated in kbp.

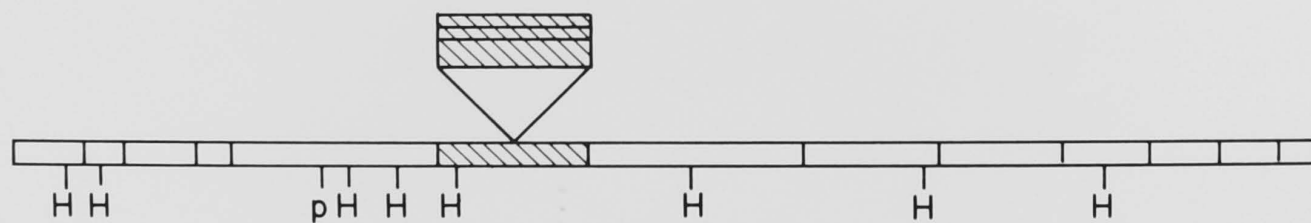
b) The proposed restriction enzyme cleavage map of the 27E2 allele. The DNA is divided into *Eco*RI fragments and nature and location of the lesions observed in the mutant are shown: ▮ the *Eco*RI fragment has been amplified. The cleavage sites of the restriction enzymes *Hind*III (H) and *Pvu*II (p) are indicated.

c) Coordinates in units of kbp.

a)



b)



c)

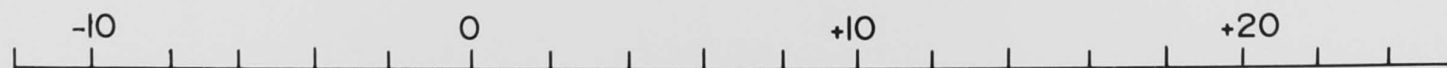




Figure 5.6: Cytology of *uncoordinated* mutant chromosomes.

Salivary gland chromosomes were prepared from third instar larvae of the following genotypes: A, S95/Canton-S; B, S86/Canton-S; C, DC803/Canton-S; D, GE230/Canton-S; E, W5/Canton-S; F, 27E2/Canton-s; G, 16.3.212/Canton-S. Individual chromosome complements were photographed at a magnification of 100X.

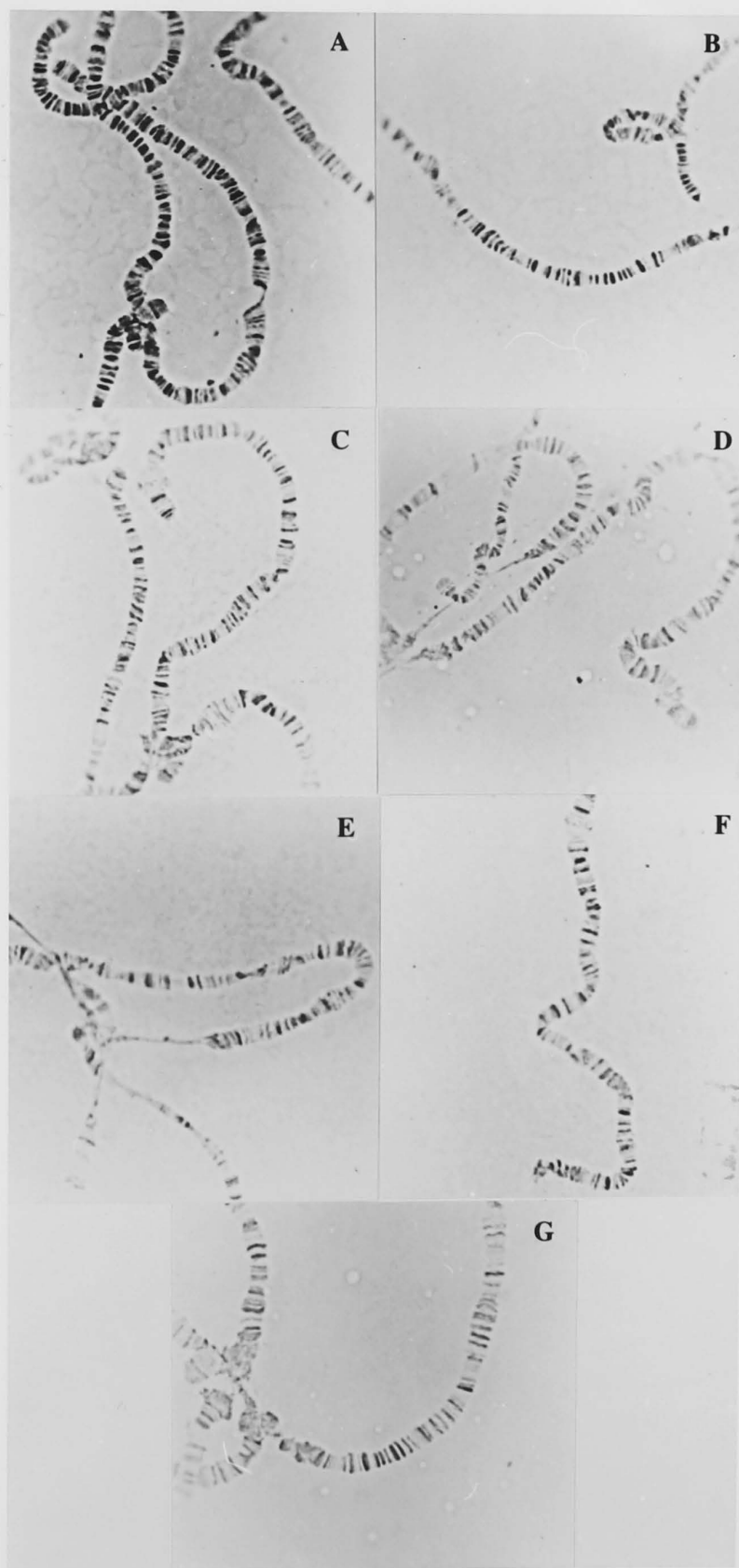
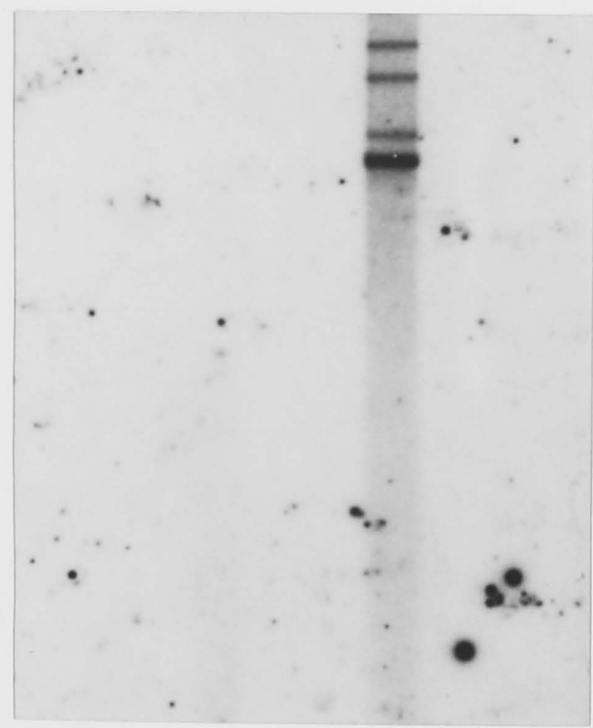


Figure 5.7(a): Identification of transcripts in *uncoordinated* mutants.

Hybridization of radiolabelled DNA from p(PP95)-P1 (left panel) to Northern blots carrying total cellular polyadenylated RNA extracted from the following genotypes: lane 1, GE230/Y; lane 2, DC803/Y; lane 4, Canton-S; lane 5, S86/S86; lane 6, S95/S95; lane 7, W5/W5; lane 9, 27E2/27E2. The 16.3.212/16.3.212 genotype is illustrated in Figure 4.11. Lane 3 contained molecular size markers and there was no sample in lane 8. RNA was extracted from late pupae of the *uncoordinated* mutants and Canton-S whole adults. The right panel shows hybridization to *Adh* and  $\lambda^+$  DNA to the same filter as the left panel. Arrows indicate the positions of the transcripts. Molecular size markers are indicated in kbp.

a)

1 2 3 4 5 6 7 8 9



— 9.5 —  
— 6.7 —  
— 4.3 —  
— 2.3 —  
— 2.0 —

1 2 3 4 5 6 7 8 9

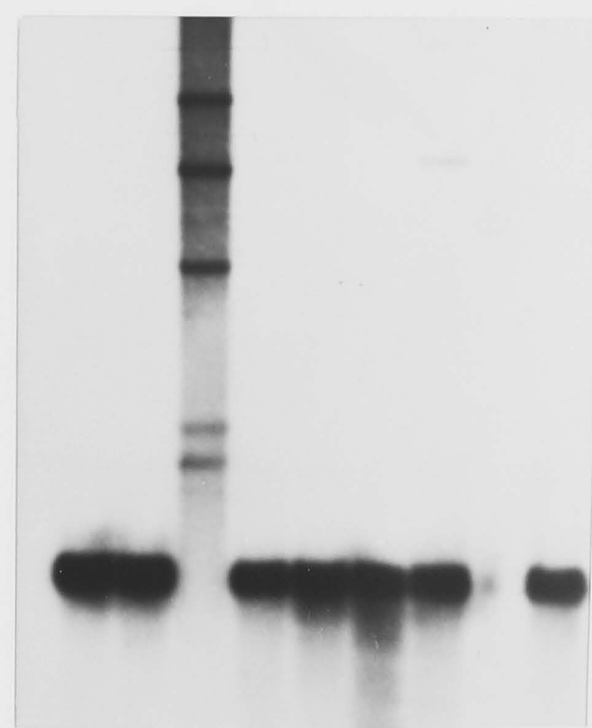


Figure 5.7(b): Identification of transcripts in *uncoordinated* mutants.

Hybridization of radiolabelled DNA from p(GM23)-C7 to Northern blots carrying total cellular RNA extracted from the same genotypes as in (a). Arrows indicate the positions of the transcripts. Molecular size markers are indicated in kbp.

b)

1 2 3 4 5 6 7 8 9



— 9.5

— 6.7

— 4.3

— 2.3

— 2.0



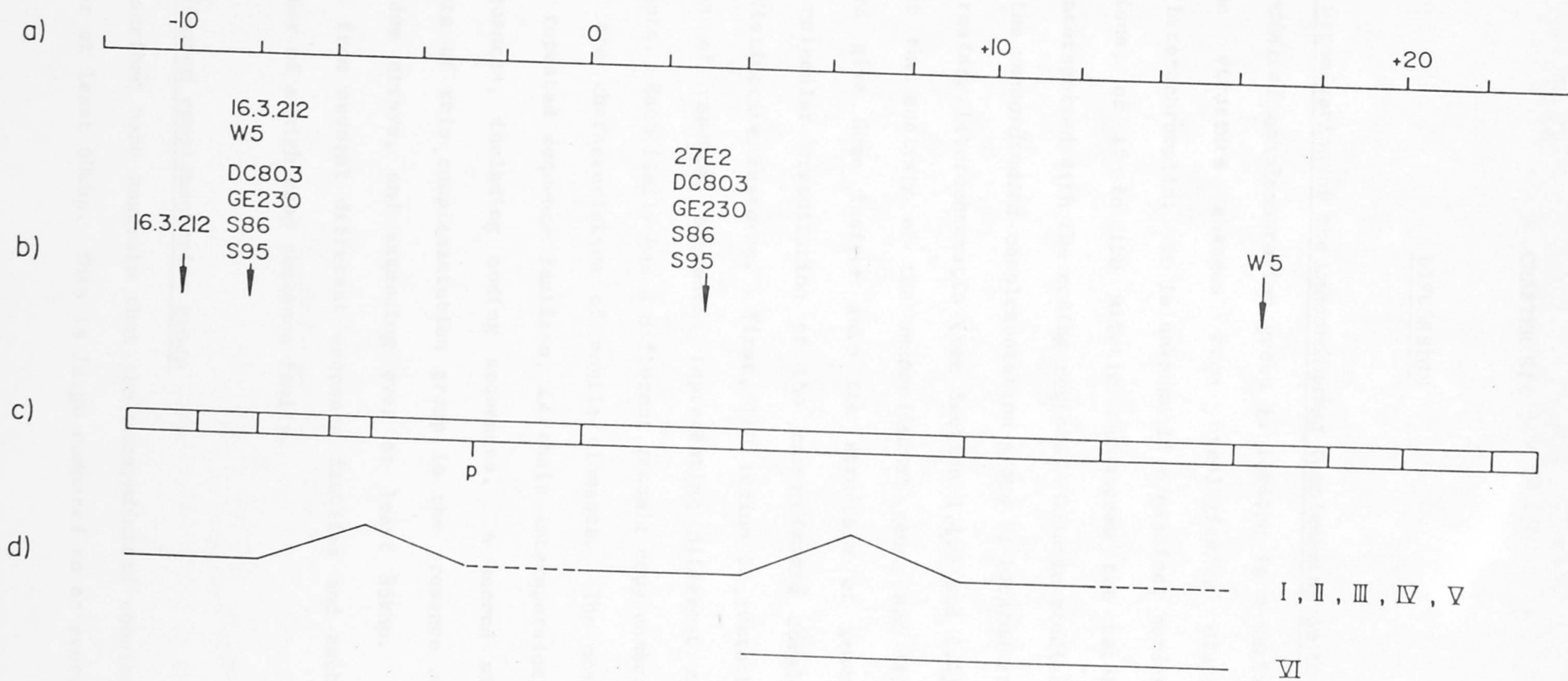
Figure 5.8: Comparison of mutagen target sites and transcript coding sequences in the *uncoordinated* gene.

a) Coordinates in units of kbp.

b) Arrows indicating the *Eco*RI fragments in which lesions occurred in each *uncoordinated* mutant allele.

c) Cloned DNA divided into *Eco*RI fragments. The cleavage site of the restriction enzyme *Pvu*II (p) is also indicated.

d) The distribution of coding regions for each mRNA transcribed from sequences localized to the *uncoordinated* complementation group. The coding —, noncoding  $\Lambda$  and possible further coding --- regions are shown. I, II, III, IV, V and VI refer to transcripts of sizes 4.35, 8.7, 9.0, 11.5, 13.0 and 7.3 kbp, respectively.



## CHAPTER SIX

DISCUSSION6.1 Molecular Organization of the *uncoordinated* Complementation Group

The *uncoordinated* complementation group is located in a region in which the chromosome structure changes from cytologically distinguishable euchromatin to heterochromatin. It is unknown if a distinct border separates these two regions, or if in the mitotic chromosome the heterochromatic sequences are interspersed with the coding regions. Genetic studies have also suggested that the *uncoordinated* complementation group is located at, or near, a site of intercalary heterochromatin (see Section 1.1.3 and 1.4). It was anticipated that the analysis of the *uncoordinated* gene, and its flanking sequences, would give some insight into the structure of genes in such regions. The molecular organization of the *uncoordinated* complementation group has two distinctive features. First, the region is characterized by several families of repeated sequences representing different classes of repetitive elements. Each family has a different genomic copy number, and may possess some of the characteristics of mobile elements. The most unusual property of the repeated sequence families, is their interspersion with low copy number sequences, including coding sequences. A second significant structural feature of this complementation group is the presence of a unit, repeated in tandem arrays, and extending over at least 35kbp. The unit contains members from several different sequence families and each unit is bounded by a member of a high copy sequence family.

6.1.1 The *uncoordinated* complementation group

The data described here suggests that the *uncoordinated* complementation group extends over at least 60kbp. This is large compared to an average value

of 32kbp per chromomere, predicted from the total DNA mass of the interband-band region. The mode value was estimated to be 13kbp in a range of 5 to about 100kbp (Rudkin, 1972; Beerman, 1972; Hogness *et al.*, 1975). Several extensive chromosomal walks have isolated stretches of genomic DNA encompassing a number of polytene chromosome bands, which have allowed an estimation of the DNA content per band. The best characterized of these is a 315kbp region encompassing approximately 13 chromomeric units from chromosome band 87D5,6 to 87E5,6 (Spierer *et al.*, 1983). The DNA sequence length per unit varies from less than 7kbp to 160kbp with most units containing 7 or 20kbp. Similar results have been obtained in other chromosomal regions, including that at chromosomes bands 3B2 to 3C2 (Pirrotta *et al.*, 1983) and 84A4,5 to 84B1,2 (Scott *et al.*, 1983). Only a small number of genes containing long stretches of DNA have been identified such as *bithorax*, *Antennapedia*, *engrailed* and *scute* (see Table 1.1). The regulation of expression of each of these genes is complex (Beachy *et al.*, 1985; Scott *et al.*, 1983; Kurer *et al.*, manuscript in prep.; Garcia-Bellido, 1979), suggesting that the *uncoordinated* complementation group may also belong to such a class of genes.

Characterization of the repeated sequence families associated with the *uncoordinated* complementation group, demonstrates that at least part of the repeated DNA consists of the type 1 insertion sequences originally identified in the ribosomal RNA genes (Glover and Hogness, 1977). Type 1 insertion sequences occur outside the nucleolus organizer in the distal centromeric heterochromatin (Dawid and Botchan, 1977; Appels and Hilliker, 1982), and *in situ* hybridization studies have suggested that these sequences occur in euchromatic regions as well (Peacock *et al.*, 1981; Cantu and Gay, 1984). Tandem arrays of type 1 insertion sequences are occasionally interspersed with other middle repetitive sequences (Kidd and Glover, 1981; Dawid *et al.*, 1981),

and Browne *et al.* (1984) have found unique sequences adjacent to type 1 insertion sequences. However, the studies reported here demonstrate a direct association between a unique coding region and this component of middle repetitive DNA.

Some of the repeated sequences from the *uncoordinated* complementation group are organized into units repeated in tandem arrays (Figure 6.1). Several mechanisms have been proposed to explain the evolution of one repeated unit into a tandem array, and that of unequal crossing over (Smith, 1976) is currently best supported experimentally (for example, Coen and Dover, 1983). The polymorphisms found in the sequences isolated from the *uncoordinated* complementation group may be a consequence of this structural organization. Unequal crossing over occurring within each strain may generate changes in the number of repeating units in the array, length heterogeneities and restriction enzyme sites.

Genes from a number of chromosomal locations have been isolated using the chromosomal walking technique, which resulted in the isolation of long stretches of genomic DNA that were essentially unique. Where repetitive elements were identified, their location was usually found to be strain specific, suggesting that they are mobile elements (Carramolino *et al.*, 1982; Kidd *et al.*, 1983; Artavanis-Tsakonas *et al.*, 1983; Segraves *et al.*, 1983; Scott *et al.*, 1983; Garber *et al.*, 1983; Bender *et al.*, 1983(a); Bender *et al.*, 1983(b); Karlik *et al.*, 1984; Kuner *et al.*, manuscript in prep.) This structural organization is in marked contrast to that of the region of the genome reported in these studies. However, a limited number of chromosomal walks have been described which share structural similarities with the *uncoordinated* complementation group. An extensive region containing repetitive elements has been isolated adjacent to the *white* locus (Pirrotta *et al.*, 1983), and although the elements have not been well characterized, at



least one component corresponds to 1.688g/cc satellite sequences (Tartof *et al.*, 1984). Sequences surrounding the *Kruppel* (Preiss *et al.*, 1985) and *hsp70* 87C (Lis *et al.*, 1978) genes and the coding region of the 93CD heat shock gene (Walldorf *et al.*, 1984) also contain repetitive elements.

It is significant that the regions so far described as containing extensive stretches of repeated elements are located at chromosomal positions genetically and cytologically defined as heterochromatic (Ananiev *et al.*, 1978). The *uncoordinated* gene is also considered to be located near, or in a region of intercalary heterochromatin. The analysis of the molecular organization of the *uncoordinated* gene therefore supports the hypothesis, proposed by Ananiev *et al.* (1978) and Belyaeva *et al.* (1984), that intercalary heterochromatin is at least partly composed of repeated sequences, and that the properties of heterochromatin may derive from the structural arrangement of repeated sequences. The organization of the repeated elements has been examined in detail for only the *uncoordinated*, *Kruppel* and *hsp70* 87C genes, and in each case, elements present in tandem arrays were identified, indicating that at least one of the requirements for heterochromatinization, as proposed by Lifschytz (1983), is a clustered structural arrangement.

Whereas all genes described so far which are surrounded by repeated elements, are located at, or near, sites of heterochromatin, a number of genes which map to such sites do not appear to be associated with repetitive elements (for example *bithorax* and *Antennapedia*). It is presently unclear as to whether sequences isolated to date from such genes do not extend through the heterochromatin. If repeated sequences are not associated with such genes, then it may not be obligatory for repetitive sequences to be present at the sites of intercalary heterochromatin as currently defined.

### 6.1.2 The euchromatin-heterochromatin transition zone

Analysis of the organization of the *uncoordinated* gene and adjacent sequences also gives some insight into the molecular organization of a euchromatin-heterochromatin transition zone. The presence of multiple repeated elements within the *uncoordinated* complementation group and other loci in regions of heterochromatin suggests that this organization may be common to genes in a zone of changing chromosome structure. This assertion is further confirmed by the analysis of a second gene in the proximal region of the X chromosome. A region of approximately 50kbp, cloned from the *little-fly-like*, or adjacent lethal complementation group, shows structural similarities to the *uncoordinated* region, in that it contains unique coding regions surrounded by members of several different repeated sequence families (Russell *et al.*, unpublished). In addition, at least 40% of the clones randomly selected from a micro-cloned library derived from subdivisions 19EF-20A contain repeated sequences (Miklos *et al.*, unpublished). In contrast, 100kbp of genomic DNA isolated from subdivision 19F6 was found to be unique (cited in Miklos *et al.*, 1984), revealing extensive heterogeneity within a small defined chromosomal region of seven bands. The detailed analysis of DNA sequences from thirteen bands in subdivisions 87DE has revealed a very uniform molecular organization (Bender *et al.*, 1983(a); Spierer *et al.*, 1983) further emphasising the unusual nature of parts of divisions 19 and 20.

The data described in this thesis is consistent with the proposal that the molecular structure of the transition zone between cytologically defined heterochromatin and euchromatin consists of unique coding regions interspersed with repetitive elements of variable lengths. It has been shown that many different families of repeated elements are present in this zone and many of the repeated sequences may be common to both centromeric and intercalary heterochromatin. Although repeated sequences have been localized to the



chromocenter, (for example, Lis *et al.*, 1978; Preiss *et al.*, 1984), the relationship between the centromeric and the chromocentral heterochromatin is unclear. The molecular organization of the *uncoordinated* complementation group therefore supports Lifschytz's (1978) model of the organization of the X chromosome transition zone deduced from genetic analysis. However, it is still not possible to determine from the molecular data if a definitive junction exists between the euchromatin and centromeric heterochromatin. Schalet and Lefevre (1976) have suggested that such a junction would be proximal to the *suppressor of forked* locus in subdivision 20F, but as yet sequences have not been characterized from this region.

## 6.2 Transcription of the *uncoordinated* Gene

### 6.2.1 Structural characteristics of the transcription unit

Analysis of the *uncoordinated* transcription unit has revealed a number of unusual features (Figure 6.1). The 5' end of the gene is probably located within the C5 fragment (coordinates +12.8 to +16.0) and there are at least three exons distributed over 28kbp of genomic DNA. Although it is unclear from the Northern blot data as to whether the C9 and C20 fragments (coordinates -7.8 to -11.0) encode exonic sequences, analysis of the two *uncoordinated* mutant alleles, W5 and 16.3.212 (Figure 5.8), suggests that coding sequences are indeed present in this region. However, as these fragments are part of a unit repeated at least three times in a tandem array, it cannot be determined how many copies of the sequence are transcribed. If all copies are transcribed, the *uncoordinated* transcription unit would be much larger, and may extend for more than 50kbp. It is also not known whether the region isolated in these studies encodes the entire *uncoordinated* gene. The identification of lesions in the genomic DNA of each of seven *uncoordinated* mutant alleles suggests that at least the majority of the coding sequences

have been isolated.\* However, the presence of the tandemly repeated sequences and the recent identification of very large genes in *Drosophila* (for example, Scott *et al.*, 1983) indicates that further confirmation is required. This may be achieved by identifying the intron-exon boundaries using the S1-nuclease mapping technique (Berk and Sharp, 1978) and characterizing cDNA clones specific to each transcript. The cloned region could also be assayed for wildtype function using the P-element transformation system (Spradling and Rubin, 1982; Rubin and Spradling, 1982). Insertion of the cloned region into the cosPneo vector (Stellar and Pirrotta, 1985) would enable the entire region to be inserted into early embryos carrying a mutation in the *uncoordinated* gene. Screening of transformants for the uncoordinated phenotype would establish if the injected sequences were sufficient to allow wildtype activity of the *uncoordinated* gene.

Repeated elements are present in the fragments between coordinates 0 to -7.8, which constitute both intronic and exonic sequences. Although repeated sequences have not been previously identified in the introns of *Drosophila* genes, they have been observed in the intronic regions of several vertebrate genes (Cochet *et al.*, 1979; Ryffel *et al.*, 1981) and are present in the heterogeneous nuclear RNA population of sea-urchins and amphibian embryos (Davidson and Posakony, 1982). The significance of such repeated sequences is not clear. However, one interpretation is that such elements have been transposed into developmentally modulated transcription units, but have no function of their own, and are evolutionarily unstable over short periods of time. An alternative view is that such elements function in coordinating gene expression, as was suggested by Davidson and Britten (1979). Several lines of evidence support this view. Firstly, the introns of some genes, such as the brain specific genes (Milner *et al.*, 1984), and the noncoding strand of the troponin-T gene (Medford *et al.*, 1984), encode a small RNA species synthesized

\* In the case of the four alleles derived from the M56i chromosome, it is unclear if the DNA sequence organisation observed is responsible for the uncoordinated mutant phenotype. The differences between the mutant and parental strain may reflect a polymorphism present in the parental population at the time of mutagenesis.

by polymerase III. Secondly, enhancer elements, which direct tissue specific expression, have been found in the introns of immunoglobulin genes (Mercola *et al.*, 1983). Finally, *Alu* sequences, known to be inserted into the intervening sequence of some genes (Ito *et al.*, 1984; Lee *et al.*, 1984), are structurally similar to the small cytoplasmic RNAs (Ullu and Tschudi, 1984), suggesting a functional role for those repetitive elements.

Short repeated sequences have been identified in the exons of several genes isolated from *Drosophila*. A sequence of approximately 180bp, termed the 'homeo-box', has been identified in a group of genes (McGinnis *et al.*, 1984; Scott and Weiner, 1984; Laughton and Scott, 1984; Poole *et al.*, 1985; Fjose *et al.*, 1985) that function in pattern formation. The identification of the homeo-box represents the first sequence to be correlated with genes that regulate a developmental function. The *opa* (Wharton *et al.*, 1985) and *suffix* (Tchurikov *et al.*, 1982) sequences have been identified in a number of transcribed sequences whose functional relationships, if any, are unclear. Thus, although a role can be assigned to some repetitive elements located in exons, the significance of the remainder is unclear. It is not currently possible to distinguish between the possibilities in relation to the *uncoordinated* gene.

The 4.35 and 7.3kbp adult transcripts were localized to the *uncoordinated* gene using RNA extracted from organisms homozygous deficient for this complementation group (Figure 4.8). The role of the 4.35kbp transcript in the functioning of the *uncoordinated* gene was further confirmed by its absence in the seven *uncoordinated* mutant alleles examined (Figure 5.7). However, in each mutant, the 7.3kbp transcript was present at wildtype levels. This questions the role of the transcript in the functioning of the *uncoordinated* gene. If the 7.3kbp transcript is derived from the *uncoordinated* transcription unit, then there are several possible explanations as to why

this transcript is unaffected by any of the mutations studied. Firstly, examination of the site of the defect in each mutant allele (Figure 5.8) demonstrates that the sequences related to the P1 fragment (coordinates 0 to +4.0) are the preferred mutagenic target sequences and the Northern blot hybridization data suggest that those sequences do not encode the 7.3kbp transcripts (Figure 4.4). Alternatively, mutant lesions may perturb the post-transcriptional processing events of the 7.3kbp species such as RNA stability, sequestering by polysomes or translation of the message. It is proposed that this also occurs in some mutations of the *white* and *rosy* loci, which have been examined at both the DNA and RNA levels (Zachar and Bingham, 1982; O'Hare *et al.*, 1984, Pirrotta and Brockl, 1984; Levis *et al.*, 1984; Covington *et al.*, 1984). Mutations caused by insertions into the large intron or the 5' flanking region of these genes, including *w<sup>i+a</sup>*, *w<sup>zm</sup>* and *ry<sup>2</sup>*, produce the wildtype transcript in the correct amounts, indicating that correct splicing, at least, can occur under these circumstances.

Another possibility is that the 7.3kbp transcript is not derived from the *uncoordinated* transcription unit, but forms part of a separate, but overlapping, transcription unit. The genomes of some viruses, such as the adenovirus group, are organized in this manner (Persson and Philipson, 1982) and more recently, overlapping transcription units have been found in *Drosophila*. Transcription units have been identified at subdivision 71CD which overlap the sequences encoding the ecdysone inducible polypeptides, EIP28 and 29 on both the same and opposite DNA strand (Cherbas, 1984). In addition, two RNA species, that overlap at the 5' ends are transcribed in opposite directions from the *Glued* locus (Swaroop *et al.*, 1985), and the gene adjacent to the ribosomal protein rp49 at subdivision 99D appears to have exonic regions in the intron of the adjacent gene (Vincent *et al.*, 1984).

The presence of the 7.3kbp transcript in higher amounts in the non-adenylated than in the poly A<sup>+</sup> class of RNA was of special interest. Very few non-adenylated transcripts have been characterized in *Drosophila* to date, although DNA-RNA tracer-driver reassociation experiments have demonstrated that only one-third of the total polysomal RNA is polyadenylated (Zimmermann *et al.*, 1980; Levy and Manning, 1981). Polysomal nonadenylated transcripts are derived from the nuclear polyadenylated RNA population, which constitutes about half the total nuclear RNA population (Zimmerman *et al.*, 1982). Non-adenylated transcripts have been identified in the RNA populations produced by only a few *Drosophila* genes, including the histones, and two genes that produce multiple transcripts, namely *bithorax* (Saint, Beachy and Hogness, manuscript in prep.) and *Antennapedia* (Scott *et al.*, 1983). Although evidence suggests that two-thirds of the polysomal RNA is non-adenylated, in most cases this RNA fraction has not been tested for transcripts. Therefore, the number of transcripts identified for each locus, thus far, may be an underestimate.

#### 6.2.2 Regulation of the *uncoordinated* transcription unit

Sequences localized to the *uncoordinated* gene appear to encode at least two, and possibly as many as six transcripts. There are several mechanisms whereby a number of transcripts can be identified by the same DNA sequence. Firstly, if a fragment identified in a particular complementation group is homologous to a transcribed sequence located elsewhere in the genome, the fragment will also hybridize to the transcript produced from that sequence. This does not seem to occur with sequences from the *uncoordinated* complementation group, since two of the transcripts were absent in organisms homozygous deficient for *uncoordinated* (Figure 4.8).



A second mechanism that generates multiple transcripts from the same DNA region requires a DNA rearrangement event which brings coding regions, separated by large distances, into close proximity. This has been observed in the production of the immunoglobulin heavy chain (Marcu, 1982), trypanosome variant glycoproteins (Borst and Cross, 1982), and in yeast mating type interconversion (Nasmyth, 1983; Beach, 1983). Examination of the *uncoordinated* gene in both the adult head, where only the 7.3kbp transcript is synthesized, and whole pupae, where both the 7.3kbp and four polyadenylated species are transcribed, failed to reveal any difference in the DNA organization. Although this observation suggests that multiple *uncoordinated* transcripts are not produced by a DNA rearrangement mechanism, it is nonetheless possible that such rearrangements occur in only a small group of cells that could not be detected in these experiments.

An alternative mechanism for generating multiple transcripts involves alternative processing of a large primary transcript. This process was initially identified for viral genes (for example, Klessig, 1977, reviewed in Ziff, 1980) and has been recently found in eukaryotes. It is proposed that multiple transcripts are generated from the *uncoordinated* gene by this process (Figure 6.1) in a temporally and spatially specific manner. The Northern blot analyses demonstrated that sequences between coordinates +9.4 and +15.0 are common to all transcripts identified as being from the *uncoordinated* gene. However, the 7.3kbp transcript carries sequences derived from a region which forms an intron in the polyadenylated transcripts, and the polyadenylated transcripts contain sequences derived from a region downstream of the coding regions of the 7.3kbp transcript. The different sizes of the polyadenylated transcripts suggest that, although these transcripts all have some exonic regions in common, the larger transcripts must contain some coding sequences not present in the smaller transcripts. It is therefore proposed that the

multiple transcripts are produced by differential selection of the 3' and internal exons, although it is not possible to determine from the available data if all the transcripts initiate from the same promotor.

In *D. melanogaster* there have been a number of reports of multiple transcripts arising from a single gene (see Table 1.1), although few of these systems have been characterized in any detail. The best defined system is the alcohol dehydrogenase gene, in which transcripts are initiated from different promoters in a temporally specific manner (Benyajati *et al.*, 1983). Differential splicing has also been observed at the 3' end of the myosin heavy chain gene (Rozek and Davidson, 1983) and the glycinamide transformylase gene (Henikoff *et al.*, 1983), whereas shuffling of internal exons is characteristic of the myosin light chain (Falkenthal *et al.*, 1985) and tropomyosin genes (Karlik *et al.*, 1984). In addition, alternative processing has been described for genes in other eukaryotic organisms in which multiple transcripts have been generated by alternative initiation sites, polyadenylation sites, 3' exons, internal exons or a combination of these (Figure 6.2). The function of multiple transcripts generated from a single gene can be assessed by examining the relationship between each transcript and its product, and in some cases, a functional relationship has been demonstrated. For example, alternative transcripts from the mouse alpha-amylase and chicken vimentin genes are synthesized in different tissues (Young *et al.*, 1981; Capetanaki *et al.*, 1983), whereas the alternative transcripts from the yeast invertase gene are translated into either secreted or intracellular protein (Carlson and Botstein, 1982). In each of these examples, the same protein is translated from each transcript. However, in other gene systems, different transcripts give rise to differing protein isoforms, or proteins with different functions, as exemplified by the rat troponin T gene (Medford *et al.*, 1984) and bovine preprotachykinin genes (Nawa *et al.*, 1984), or even different proteins. This



is illustrated by the rat calcitonin gene, which may produce calcitonin or calcitonin gene related peptide (CGRP) (Amara *et al.*, 1982). These two polypeptides differ in their C terminal amino acids, site of production and function (Rosenfeld *et al.*, 1983) and production of each is restricted either spatially or temporally. Two gene systems in *D.melanogaster* have been characterized in sufficient depth to correlate the expression of each transcript with temporal (alcohol dehydrogenase gene) or spatial compartmentalization (tropomyosin gene). The studies described here have shown that in the case of the *uncoordinated* gene, there is clear evidence of temporal, and some evidence of spatial, compartmentalization of transcripts.

The generation of alternatively processed transcripts appears to be a mechanism to ensure the correct expression of a gene whose product is required at varying times and/or locations. It can therefore be viewed as an alternative mechanism to the generation of multi-gene families for regulating gene expression. The expression of members of a multi-gene family is usually under precise regulatory controls in that each gene is transcribed in response to a different developmental cue. This is illustrated by the actin multi-gene family of *D.melanogaster*. Each of the six actin genes appears to encode a different protein isoform whose expression is correlated with muscle differentiation and reorganization (Fyrberg *et al.*, 1983). Likewise, each of the multiple RNA species encoded by a single gene is generated in response to life-stage and tissue specific signals. Studies of gene systems to date suggest that genes encoding different members of a class of proteins are not restricted to a single mechanism of gene regulation, as illustrated by the contractile protein genes of *D.melanogaster*. Both tropomyosin and actin are contractile proteins; however, two of the four described tropomyosin isoforms are generated from a single gene (Basi *et al.*, 1984). It has also been demonstrated that the temporal specificity of genes encoding the same protein

can be regulated by different mechanisms in different subgroups of the same species. This is exemplified particularly well by the alcohol dehydrogenase gene in *Drosophila*. In the *melanogaster* group, alcohol dehydrogenase is encoded by a single gene that transcribes two RNA species in a life-stage specific manner. However, in the *mojavensis* group, two genes are present, each of which produces a single protein specific to each life stage (Batterham *et al.*, 1983). At present it is not clear if particular classes of gene products are regulated by similar mechanisms or whether each coding sequence has evolved its own regulatory signals independently. Of the genes characterized from *D. melanogaster* to date (see Table 1.1), it appears that those genes coding for highly abundant structural proteins are often members of a multi-gene family, whereas the complex, homeotic and neurobiological loci are more likely to generate alternative transcripts.

### 6.3 Function of the *uncoordinated* Gene Product

The phenotype of individuals carrying mutations at the *uncoordinated* gene suggests that its product may be involved in muscle function, or in neuromuscular interactions. A large number of genes have been isolated on the basis of behavioural abnormalities, the products of which may influence the function and development of the neural and muscular systems. Classification of such genes according to the phenotype of mutant individuals, reveals classes of genes in which mutations caused paralysis in response to mechanical stress or temperature extremes, hyper- or hypo-activity or lack of coordinated movement. Genes within each of these classes may play a role in motor coordination. A number of mutants have been isolated that die as young pupae and whose movement lack coordination, and thus are phenotypically similar to *uncoordinated* individuals. These include *stoned* (Grigliatti *et al.*, 1973), *uncoordinated-like* (Schalet, 1972), *crippled* (Homyk *et al.*, 1980), *arthritic*

(Schalet and Roberts, 1973), *doomed* (Flanagan, 1977), *drop dead* (Hall, 1978) and *wobbly* (Schmidt-Nielsen and Hall, 1977). Mosaic analysis has suggested that the foci of the lesions in *doomed* and *drop dead* are the ventral nervous system and brain, respectively. However, the sites of action of the remainder have not been determined. Although the function of the product of the *uncoordinated* gene has not been determined, there are some indications, as discussed below, that it may play a role in neural tissue.

Transcriptional properties of the *uncoordinated* complementation group are not characteristic of loci in regions of heterochromatin, but share similarities with those of genes whose products have been localized to neural tissue. Firstly, the size of the transcription unit of the *uncoordinated* gene is large compared to those of other cloned genes (see Table 1.1). Although very few neurobiological genes have been isolated, the transcription units of two of the three that have been characterized (*period* and *dunce*) are large. In contrast, of the genes encoding structural components of muscle such as the actins, tropomyosins and the heavy and light chains of myosins (Table 1.1), only the transcription unit of the myosin heavy chain is larger than 4.5kbp. Of the remaining genes for which the transcription unit is large, the mutant lesion or the gene product is often associated with neural tissue; for example, *bithorax*, *Antennapedia* (Akam, 1983; Hafen *et al.*, 1984; White and Wilcox, 1984), *Notch* (Kidd *et al.*, 1983; Artavanis-Tsakonas *et al.*, 1983) and *scute* (Campuzano *et al.*, 1985). However, the significance of the products of all these genes for neural function is not clear, as their primary function may not be specific to neural tissue; for example, the *bithorax* and *Antennapedia* loci direct segment identity. It has been proposed that large transcription units are required by genes whose expression is regulated in a complex manner (Kuner *et al.*, manuscript in prep.; Beachy *et al.*, 1985). Regulation would result from the interaction of RNA species produced by the

regulatory and target loci, probably by directing the processing pathways of the primary transcript. Identification of the *bithoraxoid* unit as a regulator of *Ultrabithorax*, and as a gene which does not encode an open reading frame, supports this proposal. Kuner *et al.* (manuscript in prep.) further suggest that a gene for which the transcription unit is large, and which generates multiple transcripts, may itself encode the regulatory RNA.

The second distinctive property of the *uncoordinated* gene is the large size of its processed transcripts. Reassociation studies have suggested that the average size of polyadenylated transcripts is 1.4kbp (Izquierdo and Bishop, 1979) and detailed analyses of cloned genes (Table 1.1) have confirmed the small size of many transcripts coding for structural proteins. In contrast, transcripts produced by the *uncoordinated* gene vary in size from 4.35 to 13.5kbp, and appear to have properties in common with transcripts specific to the rat brain, in that they are larger and rarer than transcripts found in other tissues (Sutcliffe *et al.*, 1983).

Analysis of the RNA population from the adult mouse brain has revealed that it contains a highly complex class of non-adenylated transcripts that does not overlap with the poly A<sup>+</sup> population (Van Ness *et al.*, 1979; Chaudhari *et al.*, 1983). The identification of a non-adenylated transcript encoded by the *uncoordinated* complementation group, whose spatial distribution differed from that of the polyadenylated transcripts, was therefore of considerable interest. The 7.3kbp transcript was located in the adult head as well as the remainder of the body (Figure 4.5), whereas the polyadenylated transcripts were confined to the body. This result raised the interesting possibility that the 7.3kbp transcript, at least, has a role in neural function.

In summary, the properties of the *uncoordinated* transcription unit that suggest its products play a role in neural tissue include the large size of both the transcription unit and the RNA products, and the differential spatial

distribution of the transcripts. However, insufficient neurobiological genes have been isolated in *Drosophila* to assess if the characteristics of rat and mouse brain specific transcripts are also applicable to *Drosophila*. If such a relationship between the properties of the transcription unit and function of the gene product does not exist, then the alternative hypothesis, that the organization of each gene is the result of its evolutionary history, must be invoked. Although the work described in this thesis does not discriminate between neural or muscular function for the *uncoordinated* gene product, properties of the transcription unit suggest that the nervous system may be an appropriate site to initiate an investigation into the function of this gene product.

#### 6.4 Future Perspectives

##### 6.4.1 Regulation of the *uncoordinated* complementation group

The work described in this thesis clearly demonstrates that the molecular structure of the *uncoordinated* complementation group is similar to that of other regions defined cytogenetically as heterochromatic, in that repeated sequences are interspersed with coding regions and the repeated unit is organized into tandem arrays. However, properties of the transcriptional unit have many similarities to those of genes located in other regions of the genome, indicating that such properties are not involved in the differential regulation of genes in distinctive chromosomal regions. However, it is not known what effect, if any, the repeated sequences have on the mechanisms regulating the expression of the *uncoordinated* gene. If genes located in regions containing repeated sequences are subjected to regulatory mechanisms that differ from those in euchromatin, it is not known whether these differences occur at the initiation of transcription or post-transcriptionally.



In order to gain some insight into the mechanisms regulating expression of the *uncoordinated* gene, it is apparent that a number of different experimental approaches must be adopted. Firstly, it is necessary to establish whether the 7.3kbp transcript is a product of the *uncoordinated* transcription unit. The isolation and sequencing of cDNA clones and the determination of the direction of transcription of each different transcript would establish whether they have sequences and open reading frames in common. However, a more direct approach would be to employ the P-factor transformation system. Embryos deficient for the *uncoordinated* complementation group would firstly be transformed with the entire cloned region to establish if all the sequences required for wildtype function had been isolated. By undertaking a deletion analysis of the region, those sequences required to generate each transcript would be defined. Transformation of embryos using only those sequences necessary to produce the polyadenylated transcripts would determine if the 7.3kbp transcript is required to rescue the mutant phenotype. The interpretation of this experimental approach may be difficult, however, if it is not possible to prepare a construct that will only generate the poly A<sup>+</sup> transcripts.

A complementary experimental approach may be to synthesize, *in vitro*, 'anti-sense' RNA using the SP6 promotor and polymerase. Injection of wildtype embryos with the anti-sense RNA species derived from the cDNA of each transcript of the *uncoordinated* gene may inhibit the translation of the transcript. If this process generates phenocopies of the mutated gene, the effect of each transcript on development can be assessed. To date, the injection of *Drosophila* embryos and mammalian tissue culture cells has been successful. Analysis of the *uncoordinated* gene by this approach requires identification of the uncoordinated phenotype in early life-stages, or the ability to inject the anti-sense RNA into larvae or pupae.

Tissue and temporal specificity of gene expression are often determined by sequences in the immediate vicinity of the transcription unit (see section 1.3.2). Cellular components can interact with different regulatory sequences in order to activate a gene in a cell specific manner. For example, deletion analyses of the human metallothionein-II gene have defined two stretches of DNA that separately mediate the induction of the gene by heavy metal ions and by glucocorticoid hormones (Karin *et al.*, 1984). Analysis of those sequences with regulatory properties in the *uncoordinated* complementation group represents an essential step in understanding how the expression of this gene is modulated. Several techniques have been described which allow this analysis to be undertaken. Identification of sequences sensitive to DNA'ase I and S1 Nuclease would indicate regions where chromatin conformation changes, suggesting the presence of regulatory sequences. A more powerful technique would be to generate defined deletions of the flanking sequences and assay the effect of the deletion *in vivo* using the P element mediated transformation system, as described above.

It is apparent that DNA-protein interactions are an essential component of gene regulation (Section 1.3.2). Different DNA-binding proteins are associated with euchromatin and heterochromatin which are considered to be actively transcribed and non-transcribed, respectively, (see Section 1.1.2) as well as with actively transcribed and non-transcribed genes (see Section 1.3.2). Because the *uncoordinated* complementation group is located in a region of intercalary heterochromatin, it is of considerable interest to determine if the proteins associated with the adjacent repeated sequences differ from those associated with the coding regions. This would enable an assessment of the role of different protein factors, if any, in the differential expression of euchromatic and heterochromatic sequences. Tartof *et al.* (1984) have proposed a 'boundary' model to explain the formation of



heterochromatic domains (see Section 1.2.1 ). In this model, the *uncoordinated* gene would lie outside the boundaries of two heterochromatic domains, but not be heterochromatic itself. However, it is not clear if the adjacent heterochromatin would modulate expression of the coding regions. An analysis of the proteins associated with the two regions therefore represents an approach for testing this model.

Techniques have been developed for studying the interactions between DNA and proteins. Sequences which are transcribed *in vivo* can be assayed directly for activity in the presence of different nuclear extracts in an *in vitro* transcription system (Parker and Topol, 1984(a)). Such an assay also enables the factor to be followed during its purification from nuclear extracts. Analysis of those proteins that interact with sequences where transcription has not been detected, requires use of different techniques. Two dimensional fractionation of nucleosomes enabled the deduction of the composition of the nucleosome (Levinger and Varshavsky, 1982). However, the development of the band competition assay (Strauss and Varshavsky, 1984) allows the detection of specific DNA-protein interactions, particularly for DNA binding proteins with a moderate degree of sequence specificity. This technique can also be used to monitor the purification from the nuclear extract of DNA-binding proteins. The location of the DNA sequence to which the protein binds can be identified by 'foot-print' analysis (Parker and Topol, 1984(a); Wu, 1984).

Because of the importance of DNA-protein interactions in activating gene expression, and the identification of heterochromatic specific proteins, it is proposed that a comparison of the regulatory proteins represents one of the most important experimental approaches to understanding the nature of heterochromatin. Although it has been previously suggested (for example, Spofford, 1976) that the protein component of chromatin is important in determining heterochromatinization, limited progress has been made in elucidating their contribution to the formation of heterochromatin.

#### 6.4.2 Function of the *uncoordinated* gene

Although many approaches are now available for studying the function of the *uncoordinated* gene product, three types of experiments would give considerable insight into where the product is required within the organism. This would subsequently allow a detailed analysis of the function of the product. First, genetic studies, such as mosaic analysis and fate mapping (reviewed in Becker, 1976) would locate the anatomical site at which the mutant gene exerts its primary effect. Induction of gynandromorphs with patches of mutant tissue generated in different regions of the fly, would enable a correlation to be drawn between the location of cells containing the mutant gene and the *uncoordinated* phenotype. Such an analysis would locate the site of action of the *uncoordinated* gene and using the techniques of fate mapping (Hotta and Benzer, 1972) or focussing (Flanagan, 1977), it may be possible to determine if the defect causing the *uncoordinated* phenotype occurs in neural or muscular tissue. Mutants at many loci considered to be involved in motor coordination have been subjected to behavioural tests, including spontaneous motor activity, stress-induced paralysis, electroretinograms and optomotor response and courtship behaviour (for example, Homyk, 1977). The results of such tests, together with the described genetic studies, have given indications as to the site of action of the gene product. In a few instances, mutants have been subjected to detailed anatomical and physiological analyses to define precisely the nature of the lesion (for example, Wu *et al.*, 1978; Salkoff and Wyman, 1981; Ganesky and Wu, 1982; Thomas and Wyman, 1982). The *uncoordinated* gene must be subjected to analyses such as these to determine the biological significance of the gene product.





Molecular analysis of the *uncoordinated* gene would allow determination of some facets of the gene product. Sequencing of cDNA clones derived from each

transcript would enable an assessment of whether a protein product is encoded and if so, indicate the characteristics of that protein. A search of data bases containing both the amino acid and DNA sequence of known proteins would establish if the protein has been characterized previously. Perhaps the most informative experimental approach would be to localize the gene product within the cell. *In situ* hybridization of *uncoordinated* sequences to tissue sections would localize *uncoordinated* transcripts to particular tissues, and to particular cells within those tissues (Hafen *et al.*, 1983). Cloning those DNA sequences containing exons into expression vectors (Gray *et al.*, 1982) would allow the isolation of a fusion protein (Young and Davis, 1983), to which antibodies could be raised. Using the technique of immunofluorescence, the cellular location of the *uncoordinated* gene product could be visualized directly. Immuno-precipitation techniques would allow the protein product to be isolated and studied.

If the role of the *uncoordinated* gene in the organism's development is to be assessed then the function of the gene product must be determined. The available data suggests that *uncoordinated* may belong to a class of genes involved in motor coordination at the neural and/or muscular level. Although many *Drosophila* genes which may play a role in neuromuscular interactions have been isolated (reviewed in Hall, 1982), the anatomical structures influenced by those genes have been identified in only a limited number of cases, including a component of the potassium channel (Salkoff, 1983), the sodium channels (Wu and Ganetzky, 1980; Jackson *et al.*, 1984) and neurones of the giant fibre (Wyman and Thomas, 1983). It is therefore of great interest to determine if the *uncoordinated* gene is involved in the function and development of neural or muscular systems and, consequently, assess its contribution to the organism's development.

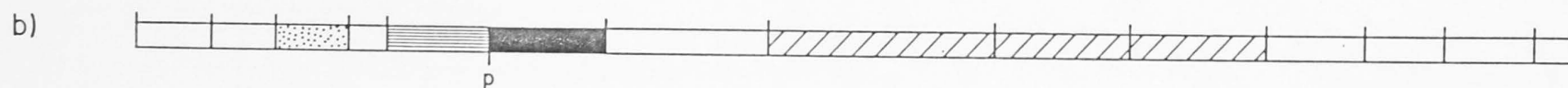
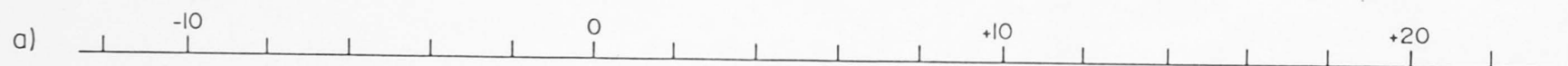
Figure 6.1: Model of transcription of the *uncoordinated* gene.

a) Coordinates in units of kbp.

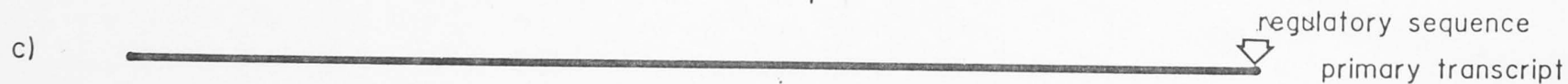
b) Cloned DNA divided into *Eco*RI fragments. The cleavage site of the restriction enzyme *Pvu*II (p) is also shown. Sequences are indicated which are highly repeated , have strong homology to multiple genomic sequences , contain short, highly abundant repeated sequences , or have weak homology to multiple genomic sequences .

c) A primary transcript synthesized from sequences in the *uncoordinated* complementation group. The location of a possible regulatory region is shown.

d) Differential processing of the primary transcript, giving rise to mRNAs of different sizes that are temporally specific. The coding —, noncoding  $\Delta$  and possible further coding --- regions are shown. The mRNA species I, II, III, IV, V and VI correspond to transcripts of 4.35, 8.7, 9.0, 11.5, 13.0 and 7.3 kbp respectively. The direction of transcription of species VI is shown.



transcription



alternative  
processing  
events

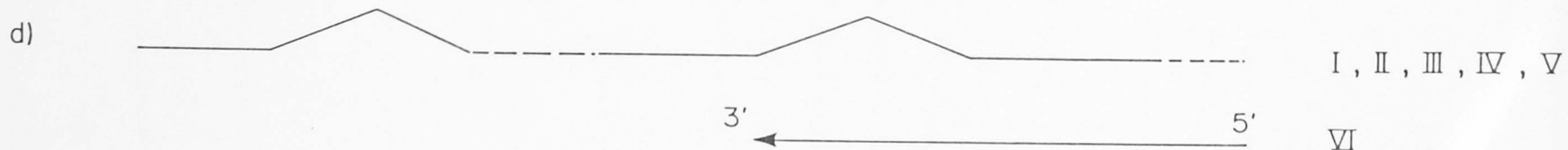


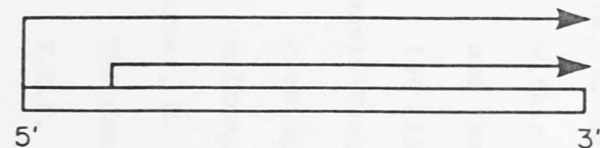
Figure 6.2: Summary of Complex Transcription Units in Higher Eukaryotes.

Multiple transcripts may be generated from a single gene by differential processing of a primary transcript. Gene systems have been described in which this occurs by selecting alternative initiation sites (1), termination sites (2), 3' exons (3) or internal exons (4).

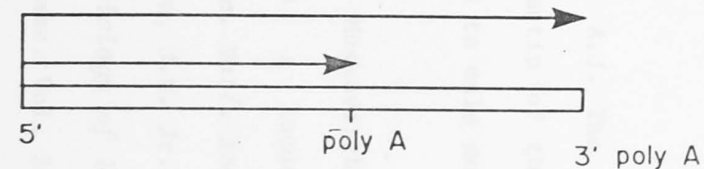


# COMPLEX TRANSCRIPTION UNITS

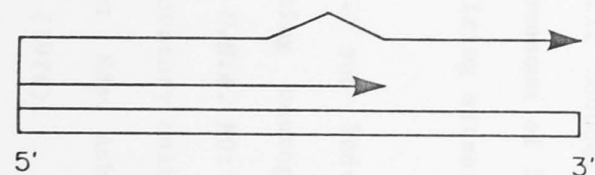
## 1. Alternative initiation sites



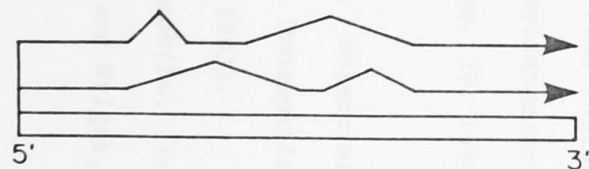
## 2. Alternative termination sites



## 3. Alternative 3' exon



## 4. Exon shuffling



## EXAMPLES

Mouse  $\alpha$ -amylase

Yeast invertase

Maize zein

Chicken myosin light chain

Chicken vimentin

Mouse  $\beta_2$ -microglobulin

Rat calcitonin

Mouse H-2K antigen

Rat troponin

Bovine preprotachykinin

Rat Fibronectin

Human Fibronectin

Mouse  $\alpha$ A-crystallin

## REFERENCE

Hagenbuchle *et al.*, 1981

Carlson and Botstein, 1982

Langridge and Feix, 1983

Nabeshima *et al.*, 1984

Zehner and Paterson, 1983

Parnes *et al.*, 1983

Amara *et al.*, 1984

Kress *et al.*, 1983

Medford *et al.*, 1984

Nawa *et al.*, 1984

Schwarzbauer *et al.*, 1983

Kornblihtt *et al.*, 1984

King and Piatigorsky, 1983

BIBLIOGRAPHY

- Akam, M.E. The location of *Ultrabithorax* transcripts in *Drosophila* tissue sections. *The EMBO Journal* 2: 2075-2084 (1983).
- Alfageme, C.R., Rudkin, G.T. and Cohen, L.H. Isolation, properties and cellular distribution of D1, a chromosomal protein of *Drosophila*. *Chromosoma* 78: 1-31 (1980).
- Allan, M., Zhu, J., Montague, P. and Paul, J. Differential response of multiple  $\epsilon$ -globin cap sites to *cis* and *trans*-acting controls. *Cell* 38: 399-407 (1984).
- Amara, S.G., Evans, R.M. and Rosenfeld, M.G. Calcitonin/Calcitonin Gene-Related Peptide transcription unit: tissue-specific expression involves selective use of alternative polyadenylation sites. *Mol. Cell. Biol.* 4: 2151-2160 (1984).
- Ananiev, E.V., Gvozdev, V.A., Ilyin, Y.V., Tchurikov, N.A. and Georgiev, G.P. Reiterated genes with varying location in intercalary heterochromatin regions of *Drosophila melanogaster* polytene chromosomes. *Chromosoma* 70: 1-17 (1978).
- Appels, R. and Hilliker, A.J. The cytogenetic boundaries of the rDNA region within heterochromatin of the X chromosome of *Drosophila melanogaster* and their relation to male meiotic pairing sites. *Genet. Res.* 39: 149-156 (1982).
- Artavanis-Tsakonas, S., Muskavitch, M.A.T. and Yedvobnick, B. Molecular cloning of *Notch*, a locus affecting neurogenesis in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 80: 1977-1981 (1983).
- Ashburner, M. and Thompson, J.N. Jr. The laboratory culture of *Drosophila*. In *The Genetics and Biology of Drosophila* Eds. Ashburner, M. and Wright, T.R.F. Academic Press. Vol. 2a: 1-109 (1978).

- Baker, W.K. Position effect variegation. *Advances in Genetics* 14: 133-169 (1968).
- Bargiello, T.A., Jackson, F.R. and Young, M.W. Restoration of circadian behavioural rhythms by gene transfer in *Drosophila*. *Nature* 312: 752-754 (1984).
- Bargiello, T.A. and Young, M.W. Molecular genetics of a biological clock in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 81: 2142-2146 (1984).
- Barnett, T., Pacht, C., Gergen, J.P. and Wensink, P.C. The isolation and characterization of *Drosophila* yolk protein genes. *Cell* 21: 729-738 (1980).
- Basi, G.S., Boardman, M. and Storti, R.V. Alternative splicing of a *Drosophila* tropomyosin gene generates muscle tropomyosin isoforms with different carboxy-terminal ends. *Mol. Cell. Biol.* 4: 2828-2836 (1984).
- Batterham, P., Lovett, J.A., Starmer, W.T. and Sullivan, D.T. Differential regulation of duplicate alcohol dehydrogenase genes in *Drosophila mojavensis*. *Develop. Biol.* 96: 346-354 (1983).
- Bautch, V.L. and Storti, R.V. Identification of a cytoplasmic tropomyosin gene linked to two muscle tropomyosin genes in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 80: 7123-7127 (1983).
- Bautch, V.L., Storti, R.V., Mischke, D. and Pardue, M.L. Organization and expression of *Drosophila* tropomyosin genes. *J. Mol. Biol.* 162: 231-250 (1982).
- Beach, D.H. Cell type switching by DNA transposition in fission yeast. *Nature* 305: 682-687 (1983).
- Beachy, P.A., Helfand, S.L. and Hogness, D.S. Segmental distribution of bithorax complex proteins during *Drosophila* development. *Nature* 313: 545-551 (1985).

- Beall, C.J. and Hirsh, J. High levels of intron-containing RNAs are associated with expression of the *Drosophila* dopa decarboxylase gene. *Mol. Cell. Biol.* 4: 1669-1674 (1984).
- Becker, H.J. Mitotic recombination. In *The Genetics and Biology of Drosophila*. Eds. Ashburner, M. and Novitski, E. Academic Press. Vol. 1c: 1019-1087 (1976).
- Beermann, W. Chromomeres and Genes. In *Results and Problems in Cell Differentiation*. Ed. Beermann, W. Springer-Verlag. Vol. 4: 1-33 (1972).
- Belyaeva, E. S., Ananiev, E.V. and Gvozdev, V.A. Distribution of mobile dispersed genes (*mdg-1* and *mdg-3*) in the chromosomes of *Drosophila melanogaster*. *Chromosoma* 90: 16-19 (1984).
- Bender, W., Akam, M., Karch, F., Beachy, P.A., Peifer, M., Spierer, P., Lewis, E.B. and Hogness, D.S. Molecular genetics of the bithorax complex in *Drosophila melanogaster*. *Science* 221: 23-29 (1983(a)).
- Bender, W., Spierer, P. and Hogness, D.S. Chromosomal walking and jumping to isolate DNA from the *ace* and *rosy* loci and the bithorax complex in *Drosophila melanogaster*. *J. Mol. Biol.* 168: 17-33 (1983(b)).
- Benton, W.D. and Davis, R.W. Screening  $\lambda$ gt recombinant clones by hybridization to single plaques *in situ*. *Science* 196: 180-182 (1977).
- Benyajati, C., Spoerel, N., Haymerle, H. and Ashburner, M. The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5' end in different developmental stages. *Cell* 33: 125-133 (1983).
- Berg, R., Engels, W.R. and Kramer, R.A. Site specific X-chromosome rearrangements from hybrid dysgenesis in *Drosophila melanogaster*. *Science* 210: 427-429 (1980).
- Berger, E.M. The regulation and function of small heat-shock protein synthesis. *Develop. Genet.* 4: 255-265 (1984).
- Berk, A.J. and Sharp, P.A. Spliced early mRNAs of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 75: 1274-1278 (1978).

- Bialojan, S., Falkenburg, D. and Renkawitz-Pohl, R. Characterization and developmental expression of  $\beta$ -tubulin genes in *Drosophila melanogaster*. *The EMBO Journal* 3: 2543-2548 (1984).
- Bingham, P.M., Kidwell, M.G. and Rubin, G.M. The molecular basis of P-M hybrid dysgenesis: the role of the P element, a P-strain-specific transposon family. *Cell* 29: 995-1004 (1984).
- Blattner, F.R., Williams, B.G., Brechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L., Greenwald, D.J., Kiefer, D.O., Moore, D.O., Schumm, J.W., Sheldon, E.L. and Smithies, O. Charon phages: safer derivatives of bacteriophage lambda for DNA cloning. *Science* 196: 161-169 (1977).
- Bodenstein, D. The post-embryonic development of *Drosophila*. In *Biology of Drosophila* Ed. Demerec, M. pp275-367 John Wiley and Sons (1950).
- Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2: 95-113 (1977).
- Borst, P. and Cross, G.A.H. Molecular basis for *Trypanosome* antigenic variation. *Cell* 29: 291-303 (1982).
- Bourouis, M. and Richards, G. Remote regulatory sequences of the *Drosophila* *Sgs-3* as revealed by P-element transformation. *Cell* 40: 349-357 (1985).
- Brown, D.D. The role of stable complexes that repress and activate eukaryotic genes. *Cell* 37: 359-365 (1984).
- Brown, D.D. and Dawid, I.B. Specific gene amplification in oocytes. *Science* 160: 272-280 (1968).
- Brown, S.W. Heterochromatin. *Science* 151: 417-425 (1966).
- Browne, M.J., Read, C.A., Roiha, H. and Glover, D.M. Site specific insertion of a type 1 rDNA element into a unique sequence in the *Drosophila melanogaster* genome. *Nucleic Acids Res.* 12: 9111-9122 (1984).

- Brutlag, D.L. Molecular arrangement and evolution of heterochromatin. *Ann. Rev. Genetics* 14: 121-144 (1980).
- Brutlag, D., Carlson, M., Fry, K. and Hsieh, T.S. DNA sequence organization in *Drosophila* heterochromatin. *Cold Spring Harbor Symp. Quant. Biol.* 42: 1137-1146 (1977).
- Bucheton, A., Paro, R., Sang, H.M., Pelisson, A. and Finnegan, D.J. The molecular basis of I-R hybrid dysgenesis in *Drosophila melanogaster*. Identification, cloning and properties of the I factor. *Cell* 38: 153-163 (1984).
- Bulthoff, H. Isolation of sex-linked mutants disturbed in visual orientation. *Drosophila Information Service* 58: 32-33 (1982).
- Burns, D.K., Stark, B.C., Macklin, M.D. and Chooi, W.Y. Isolation and characterization of cloned DNA sequences containing protein genes of *Drosophila melanogaster*. *Mol. Cell. Biol.* 4: 2643-2652 (1984).
- Campbell, D.A., Thornton, D.A. and Boothroyd, J.C. Apparent discontinuous transcription of *Trypanosoma brucei* variant surface antigen genes. *Nature* 311: 350-355 (1984).
- Campuzano, S., Carromolino, L., Cabrera, C.V., Ruiz-Gomez, M., Villares, R., Boronat, A. and Modolell, J. Molecular genetics of the *achaete-scute* gene complex of *Drosophila melanogaster*. *Cell* 40: 327-338 (1985).
- Cantu, E.S. and Gay, H. Localization of chromosomal DNA sequences homologous to ribosomal gene type 1 insertion DNA in *Drosophila melanogaster*. *Mol. Gen. Genet.* 196: 345-349 (1984).
- Capetanaki, Y.G., Ngai, J., Flytzanis, C.N. and Lazarides, E. Tissue-specific expression of two mRNA species transcribed from a single vimentin gene. *Cell* 35: 411-420 (1983).
- Carlson, M. and Botstein, D. Two differentially regulated mRNAs with different 5' ends encode secreted and intracellular forms of yeast invertase. *Cell* 28: 145-154 (1982).



- Carlson, M. and Brutlag, D. Cloning and characterization of a complex satellite DNA from *Drosophila melanogaster*. *Cell* 11: 371-381 (1977).
- Carlson, M. and Brutlag, D. A gene adjacent to satellite DNA in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 75: 5898-5902 (1978(a)).
- Carlson, M. and Brutlag, D. One of the copia genes adjacent to satellite DNA in *Drosophila melanogaster*. *Cell* 15: 733-742 (1978(b)).
- Carlson, M. and Brutlag, D. Different regions of a complex satellite DNA vary in size and sequence of the repeating unit. *J. Mol. Biol.* 135: 483-500 (1979).
- Carmichael, G.G. and McMaster, G.K. The analysis of nucleic acids in gels using glyoxal and acridine orange. *Methods in Enzymology* 65: 380-391 (1980).
- Carramolino, L., Ruiz-Gomez, M., Guerrero, M., Campuzano, S. and Modolell, J. DNA map of mutations at the *scute* locus of *Drosophila melanogaster*. *The EMBO Journal* 1: 1185-1193 (1982).
- Chang, A.C.Y. and Cohen, S.N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* 134: 1141-1156 (1978).
- Charney, P., Treisman, R., Mellon, P., Choa, M., Axel, R. and Maniatis, T. Differences in human  $\alpha$ - and  $\beta$ -globin gene expression in mouse erythroleukemia cells: the role of intragenic sequences. *Cell* 38: 251-263 (1984).
- Chaudhari, N. and Hahn, W.E. Genetic expression in the developing brain. *Science* 220: 924-928 (1983).
- Cherbas, P. Transcription at the EIP locus. Crete Workshop (1984).
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18: 5294-5299 (1979).

- Chovnick, A., McCarron, M., Hilliker, A., O'Donnell, J., Gelbart, W. and Clark, S. Gene organization in *Drosophila*. *Cold Springs Harbor Symp. Quant. Biol.* 42: 1011-1021 (1977).
- de Cicco, D.V. and Spradling, A.C. Localization of a *cis*-acting element responsible for the developmentally regulated amplification of *Drosophila* chorion genes. *Cell* 38: 45-54 (1984).
- Clark, S.H., Daniels, S., Rushlow, C.A., Hilliker, A.J. and Chovnick, A. Tissue-specific and pre-translational character of variants of the *rosy* locus control element in *Drosophila melanogaster*. *Genetics* 108: 953-968 (1984).
- Cochet, M., Gannon, F., Hen, R., Maroteaux, L., Perrin, F. and Chambon, P. Organization and sequence studies of the 17-piece chicken conalbumin gene. *Nature* 282: 567-574 (1979).
- Coen, E.S., and Dover, G. Unequal exchanges and the coevolution of X and Y rDNA arrays in *Drosophila melanogaster*. *Cell* 33: 849-855 (1983).
- Cohen, R.S. and Meselson, M., Inducible transcription and puffing in *Drosophila melanogaster* transformed with *hsp 70*-phage  $\lambda$  hybrid heat shock genes. *Proc. Natl. Acad. Sci. U.S.A.* 81: 5509-5513 (1984).
- Cooper, K.W. Cytogenetic analysis of major heterochromatic elements (especially Xh and Y) in *Drosophila melanogaster* and the theory of heterochromatin. *Chromosoma* 10: 535-588 (1959).
- Corces, V., Pellicer, A., Axel, R. and Meselson, M. Integration, transcription and control of a *Drosophila* heat shock gene in mouse cells. *Proc. Natl. Acad. Sci. U.S.A.* 78: 7038-7042 (1981).
- Covington, M., Fleenor, D. and Devlin, R.B. Analysis of xanthine dehydrogenase mRNA levels in mutants affecting the expression of the *rosy* locus. *Nucleic Acids Res.* 12: 4559-4573 (1984).

- Crowley, T.E., Bond, M.W. and Meyerowitz, E.M. The structural genes for three *Drosophila* glue proteins reside at a single polytene chromosome puff locus. *Mol. Cell. Biol.* 3: 623-634 (1983).
- Crowley, T.E., Mathers, P.H. and Meyerowitz, E.M. A *trans*-acting regulatory product necessary for expression of the *Drosophila melanogaster* 68C glue gene cluster. *Cell* 39: 149-156 (1984).
- Darnell, J.E. Jr. Variety in the level of gene control in eukaryotic cells. *Nature* 297: 365-371 (1982).
- Davidson, E.H. and Britten, R.J. Regulation of gene expression: possible role of repetitive sequences. *Science* 204: 1052-1059 (1979).
- Davidson, E.H. and Posakony, J.W. Repetitive sequence transcripts in development. *Nature* 297: 633-635 (1982).
- Davis, R.L. and Davidson, N. Isolation of the *Drosophila melanogaster dunce* chromosomal region and recombinational mapping of *dunce* sequences with restriction site polymorphisms as genetic markers. *Mol. Cell. Biol.* 4: 358-367 (1984).
- Davis, R.W., Bottstein, D. and Roth, J.R. *A manual for genetic engineering: advanced bacterial genetics*. Cold Spring Harbor Laboratory. New York. (1982).
- Davison, B.L., Egly, J.M., Mulvihill, E.R. and Chambon, P. Formation of stable pre-initiation complexes between eukaryotic class B transcription factors and promotor sequences. *Nature* 301: 680-686 (1983).
- Dawid, I.B. and Botchan, P. Sequences homologous to ribosomal insertions occur in the *Drosophila* genome outside the nucleolus organizer. *Proc. Natl. Acad. Sci. U.S.A.* 74: 4233-4237 (1977).
- Dawid, I.B., Long, E.O. and di Nocera, P. Ribosomal insertion-like elements in *Drosophila melanogaster* are interspersed with mobile sequences. *Cell* 25: 399-408 (1981).

- Dawid, I.B., Weillauer, P.K. and Long, E.O. Ribosomal DNA in *Drosophila melanogaster*. I. Isolation and characterisation of cloned fragments. *J. Mol. Biol.* 126: 749-768 (1978).
- Deak, I.I., Bellamy, P.R., Buenz, M., Bubijs, Y., Fenner, E., Gollin, M., Rahmi, A., Ramp, T., Reinhardt, C.A. and Cotton, B. Mutations affecting the indirect flight muscles of *Drosophila melanogaster*. *J. Embryol. Exp. Morph.* 69: 61-81 (1982).
- Denhardt, D.T. A membrane filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* 23: 641-652 (1966).
- Dudai, Y., Jan, Y.N., Byers, D., Quinn, W.G. and Benzer, S. *dunce*, a mutant of *Drosophila* deficient in learning. *Proc. Natl. Acad. Sci. U.S.A.* 73: 1684-1688 (1976).
- Dudler, R. and Travers, A.A. Upstream elements necessary for optimal function of the *hsp 70* promoter in transformed flies. *Cell* 38: 391-398 (1984).
- Dunsmuir, P., Brorein, W.J. Jr, Simon, M.A. and Rubin, G.M. Insertion of the *Drosophila* transposable element *copia* generates a 5 base pair duplication. *Cell* 21: 575-579 (1980).
- Evans, T., Schon, E., Gora-Maslak, G., Patterson, J. and Efstratiadis, A.F. SI-hypersensitive sites in eukaryotic promoter regions. *Nucleic Acids Res.* 12: 8043-8058 (1984).
- Falkenthal, S., Parker, V.P. and Davidson, N. Developmental variations in the splicing pattern of transcripts from the *Drosophila* gene encoding myosin alkali light chain result in different carboxyl-terminal amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A.* 82: 449-453 (1985).
- Falkenthal, S., Parker, V.P., Mattox, W.W. and Davidson, N. *Drosophila melanogaster* has only one myosin alkali light-chain gene which encodes a protein of considerable amino acid sequence homology to chicken myosin alkali light chains. *Mol. Cell. Biol.* 4: 956-965 (1984).

- Finnegan, D.J., Rubin, G.M., Young, M.W. and Hogness, D.S. Repeated gene families in *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.* 42: 1053-1963 (1977).
- Fjose, A., McGinnis, W.J. and Gehring, W.J. Isolation of a homoeo box containing gene from the *engrailed* region of *Drosophila* and the spatial distribution of its transcripts. *Nature* 313: 284-289 (1985).
- Fjose, A., Polito, L.C., Weber, U. and Gehring, W.J. Developmental expression of the *white* locus of *Drosophila melanogaster*. *The EMBO Journal* 3: 2087-2094 (1984).
- Flanagan, J.R. A method for fate mapping the foci of lethal and behavioural mutants in *Drosophila melanogaster*. *Genetics* 85: 587-607 (1977).
- Frischauf, A.M., Lehrach H., Poustka, A. and Murray, N. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170: 827-842 (1983).
- Fruscoloni, P., Al-Atia, G.R. and Jacobs-Lorena, M. Translational regulation of a specific gene during oogenesis and embryogenesis of *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 80: 3359-3363 (1983).
- Fyrberg, E.A., Bond, B.J., Hershey, N.D., Mixter, K.S. and Davidson, N. The actin genes of *Drosophila*: protein coding regions are highly conserved but intron positions are not. *Cell* 24: 107-116 (1981).
- Fyrberg, E.A., Kindle, K.L., Davidson, N. and Sodja, A. The actin genes of *Drosophila*: a dispersed multigene family. *Cell* 19: 365-378 (1980).
- Fyrberg, E.A., Mahaffey, J.W., Bond, B.J. and Davidson, N. Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. *Cell* 33: 115-123 (1983).
- Gall, J.G. Repetitive DNA in *Drosophila*. In *Molecular Cytogenetics* Eds. Hamkalo, B.A. and Papaconstantinou, J. Plenum Press. pp59-74 (1973).



- Ganetzky, B. and Wu, C.F. Indirect suppression involving behavioural mutants with altered nerve excitability in *Drosophila melanogaster*. *Genetics* 100: 597-614 (1982).
- Garabedian, M.J., Hung, M.C. and Wensink, P.C. Independent control elements that determine yolk protein gene expression in alternative *Drosophila* tissues. *Proc. Natl. Acad. Sci. U.S.A.* 82: 1396-1400 (1985).
- Garber, R.L., Kuroiwa, A. and Gehring, W.J. Genomic and cDNA clones of the homeotic locus *Antennapedia* in *Drosophila*. *The EMBO Journal* 2: 2027-2036 (1983).
- Garcia-Bellido, A. Genetic analysis of the *achaete-scute* system of *Drosophila melanogaster*. *Genetics* 91: 491-520 (1979).
- Garcia-Bellido, A. and Santamaria, P. Developmental analysis of the *achaete-scute* system of *Drosophila melanogaster*. *Genetics* 88: 469-486 (1978).
- Gartler, S.M. and Riggs, A.D. Mammalian X-chromosome inactivation. *Ann. Rev. Genet.* 17: 155-190 (1983).
- Gatti, M., Smith, D.A. and Baker, B.S. A gene controlling condensation of heterochromatin in *Drosophila melanogaster*. *Science* 221: 83-85 (1983).
- Gehring, W.J. The homeotic box: A key to the understanding of development. *Cell* 40: 3-5 (1985).
- Gehring, W.J., Klemenz, R., Weber, U. and Kloter, U. Functional analysis of the *white*<sup>+</sup> gene of *Drosophila* by P-factor-mediated transformation. *The EMBO Journal* 3: 2077-2085 (1984).
- Gilbert, W., Majors, J. and Mascam, A. How proteins recognize DNA sequences. In *Dahlem Workshop on Organization and Expression of Chromosomes*. Eds. Allfrey, V.G., Bautz, E.K.F., McCathy, B.J., SchMike, R.T. and Tissieres, A. Akakon Verlagsgesellschaft. pp 167-178 (1976).
- Glover, D.M. and Hogness, D.S. A novel arrangement of the 18S and 28S sequences in a repeating unit of *Drosophila melanogaster* rDNA. *Cell* 10: 167-176 (1977).



- Goldberg, D.A. Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. *Proc Natl. Acad. Sci. U.S.A.* 77: 5794-5798 (1980).
- Goldberg, D.A., Posakony, J.W. and Maniatis, T. Correct developmental expression of a cloned alcohol dehydrogenase gene transduced into the *Drosophila* germ line. *Cell* 34: 59-73 (1983).
- Gray, M.R., Colet, H.V., Guarente, L. and Rosbash, M. Open reading frame cloning: identification, cloning and expression of open reading frame DNA. *Proc. Natl. Acad. Sci. U.S.A.* 79: 6598-6602 (1982).
- Green, M.M. and Miklos, G.L.G. The generation of deleted X chromosomes using the male recombination (MR) system. *Drosophila Information Service* 59: 45-46 (1983).
- Griffin-Shea, R., Thireos, G. and Kafatos, F.C. Organisation of a cluster of four chorion genes in *Drosophila* and its relationship to developmental expression and amplification. *Dev. Biol.* 91: 325-336 (1982).
- Griffin-Shea, R., Thireos, G., Kafatos, F.C., Petri, W.H. and Villa-Komaroff, L. Chorion cDNA clones of *D.melanogaster* and their use in studies of sequence homology and chromosomal location of chorion genes. *Cell* 19: 915-922 (1980).
- Grigliatti, T.A., Hall, L., Rosenbluth, R. and Suzuki, D.T. Temperature-sensitive mutations in *Drosophila melanogaster*. XIV. A selection of immobile adults. *Mol. Gen. Genet.* 120: 107-114 (1973).
- Guild, G.M. Molecular analysis of a developmentally regulated gene which is expressed in the larval salivary gland of *Drosophila*. *Devel. Biol.* 102: 462-470 (1984).
- Guild, G.M. and Shore, E.M. Larval salivary gland secretion proteins in *Drosophila*. Identification and characterization of the *Sgs-5* structural gene. *J. Mol. Biol.* 179: 289-314 (1984).

- Gvozdev, V.A., Ananiev, E.V., Katelyanskaya, A.E. and Zhimulev, I.F. Transcription of intercalary heterochromatin regions in *Drosophila melanogaster* cell culture. *Chromosoma* 80: 177-190 (1980).
- Hafen, E., Levine, M., Garber, R.L. and Gehring, W.J. An improved 'in situ' hybridization method for the detection of cellular RNAs in *Drosophila* tissue sections and its application for localising transcripts of the homeotic *Antennapedia* gene complex. *The EMBO Journal* 2: 617-623 (1983).
- Hafen, E., Levine, M. and Gehring, W.J. Regulation of *Antennapedia* transcript distribution by the *bithorax* complex in *Drosophila*. *Nature* 307: 287-289 (1984).
- Hagenbuchle, O., Tosi, M., Schibler, V., Bovey, R., Wellauer, P.K. and Young, R.A. Mouse liver and salivary gland  $\alpha$ -amylase mRNAs differ only in 5' non-translated sequences. *Nature* 289: 643-646 (1981).
- Hall, J.C. Genetics of the nervous system in *Drosophila*. *Quart. Rev. Biophys.* 15: 223-479 (1982).
- Hall, J.C. Complex brain and behavioural functions disrupted by mutations in *Drosophila*. *Develop. Genet.* 4: 355-378 (1984).
- Hannah, A. Localization and function of heterochromatin in *Drosophila melanogaster*. *Advances in Genetics* 4: 87-125 (1951).
- Harshman, L.G. A technique for the production of *Drosophila* salivary gland chromosomes. *Drosophila Information Service* 52: 164 (1977).
- Hazelrigg, T., Levis, R. and Rubin, G.M. Transformation of *white* locus DNA in *Drosophila*: dosage compensation, *zeste* interaction, and position effects. *Cell* 36: 469-481 (1984).
- Henikoff, S. Position effects and variegation enhancers in an autosomal region of *Drosophila melanogaster*. *Genetics* 93: 105-115 (1979).

- Henikoff, S. Position effect variegation and chromosome structure of a heat shock puff in *Drosophila*. *Chromosoma* 83: 381-393 (1981).
- Henikoff, S., Sloan, J.S. and Kelly, J.D. A *Drosophila* metabolic gene transcript is alternatively processed. *Cell* 34: 405-414 (1983).
- Herbert, E. and Uhler, M. Biosynthesis of polyprotein precursors to regulatory peptides. *Cell* 30: 1-2 (1982).
- Hilliker, A.J. Genetic analysis of the centromeric heterochromatin of chromosome 2 of *Drosophila melanogaster*: deficiency mapping of EMS-induced lethal complementation groups. *Genetics* 88: 765-782 (1976).
- Hilliker, A.J. and Appels, R. Pleiotropic effects associated with the deletion of heterochromatin surrounding rDNA on the X chromosome of *Drosophila*. *Chromosoma* 86: 469-490 (1982).
- Hilliker, A.J., Appels, R. and Schalet, A. The genetic analysis of *D. melanogaster* heterochromatin. *Cell* 21: 607-619 (1980).
- Hilliker, A.J. and Holm, D.G. Genetic analysis of the proximal region of chromosome 2 of *Drosophila melanogaster*. I. Detachment products of compound autosomes. *Genetics* 81: 705-721 (1975).
- Hirsh, J. and Davidson, N. Isolation and characterization of the dopa decarboxylase gene of *Drosophila melanogaster*. *Mol. Cell. Biol.* 1: 475-485 (1981).
- Hirsh, J., Morgan, B. and Scholnick, S. Delimiting essential sequences of the dopa decarboxylase gene. 25<sup>th</sup> Annual *Drosophila Research Conference* (1984).
- Hogness, D.S., Wensink, P.C., Glover, D.M., White, R.L., Finnegan, D.J. and Donaldson, J.E. The arrangement of DNA sequences in the chromosomes of *Drosophila melanogaster*. In *The Eukaryote Chromosome*. ANU Press. pp. 1-33 (1975).

- Holmes, D.S. and Quigley, M. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* 114: 193-197 (1981).
- Holmgren, R., Livak, K., Morimoto, R., Freund, R. and Meselson, M. Studies of cloned sequences from four *Drosophila* heat shock loci. *Cell* 18: 1359-1370 (1979).
- Holmquist, G. Hoechst 33258 fluorescent staining of *Drosophila* chromosomes. *Chromosoma* 49: 333-356 (1975).
- Homyk, T. Jr. Behavioural mutants of *Drosophila melanogaster*. II. Behavioural analysis and focus mapping. *Genetics* 87: 105-128 (1977).
- Homyk, T. Jr. and Sheppard, D.E. Behavioural mutants of *Drosophila melanogaster*. I. Isolation and mapping of mutations which decrease flight ability. *Genetics* 87: 95-104 (1977).
- Homyk, T. Jr., Szidonya, J. and Suzuki, D.T. Behavioural mutants of *Drosophila melanogaster*. III. Isolation and mapping of mutations by direct visual observations of behavioural phenotypes. *Mol. Gen. Genet.* 177: 553-565 (1980).
- Hotta, Y. and Benzer, S. Mapping of behaviour in *Drosophila* mosaics. *Nature* 240: 527-535 (1972).
- Hsieh, T.S. and Brutlag, D.L. A protein that preferentially binds *Drosophila* satellite DNA. *Proc. Natl. Acad. Sci. U.S.A.* 76: 726-730 (1979(a)).
- Hsieh, T.S. and Brutlag, D.L. Sequence and sequence variation within the 1688g/cm<sup>3</sup> satellite DNA of *Drosophila melanogaster*. *J. Mol. Biol.* 135: 465-481 (1979(b)).
- Ikenaga, H. and Saigo, K. Insertion of a movable genetic element, 297, into the TATA box of the H3 histone gene in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 79: 4143-4147 (1982).

- Ilyin, Y.V., Chmeliaukaite, V.G., Ananiev, E.V. and Georgiev, G.P. Isolation and characterization of a new family of mobile dispersed genetic elements, *mdg3*, in *Drosophila melanogaster*. *Chromosoma* 81: 27-53 (1980).
- Ingolia, T.D. and Craig, E.A. *Drosophila* gene related to the major heat shock-induced gene is transcribed at normal temperatures and not induced by heat shock. *Proc. Natl. Acad. Sci. U.S.A.* 79: 525-529 (1982).
- Inouye, S., Yuki, S. and Saigo, K. Sequence-specific insertion of the *Drosophila* transposable genetic element *17.6*. *Nature* 310: 332-333 (1984).
- Ito, R., Sato, K., Helmer, T., Jay, G. and Agarwal, K. Structural analysis of the gene encoding human gastrin: the large intron contains an *Alu* sequence. *Proc. Natl. Acad. Sci. U.S.A.* 81: 4662-4666 (1984).
- Izant, J.G. and Weintraub, H. Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell* 36: 1007-1015 (1984).
- Izquierdo, M. and Bishop, J.O. An analysis of cytoplasmic RNA populations in *Drosophila melanogaster*, Oregon R. *Biochem. Gen.* 17: 473-497 (1979).
- Jackson, F.R., Wilson, S.D., Strichartz, G.R. and Hall, L.M. Two types of mutants affecting voltage-sensitive sodium channels in *Drosophila melanogaster*. *Nature* 308: 189-191 (1984).
- John, B. and Miklos, G.L.G. Functional aspects of satellite DNA and heterochromatin. *Int. Rev. Cytol.* 58: 1-113 (1979).
- Judd, B.H. Genetic units of *Drosophila* - Complex loci. In *The Genetics and Biology of Drosophila*. Eds. Ashburner, M. and Novitski, E. Academic Press. Vol. 1b: 767-799 (1976).
- Kalfayan, L. and Wensink, P.C. Developmental regulation of *Drosophila*  $\alpha$ -tubulin genes. *Cell* 29: 91-98 (1982).

- Karin, M., Haslinger, A., Holtgreve, H., Richards, R.L., Kerouter, P., Westphal, H.N. and Beato, M. Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionin-IIA gene. *Nature* 308: 513-519 (1984).
- Karlik, C.C., Mahaffey, J.W., Coutu, M.D. and Fyrberg, E.A. Organization of contractile protein genes within the 88F subdivision of the *D. melanogaster* third chromosome. *Cell* 37: 469-481 (1984).
- Kaufman, B.P. Somatic mitosis of *Drosophila melanogaster*. *J. Morphol.* 56: 125-155 (1934).
- Keimar, C.C. and Novick, R.P. Plasmid pT181 replication is regulated by two counter-transcripts. *Proc. Natl. Acad. Sci. U.S.A.* 82: 638-642 (1985).
- Kidd, S.J. and Glover, D.M. A DNA segment from *D. melanogaster* which contains five tandemly repeating units homologous to the major rDNA insertion. *Cell* 19: 103-119 (1980).
- Kidd, S., Lockett, T.J. and Young, M.W. The *Notch* locus of *Drosophila melanogaster*. *Cell* 34: 421-433 (1983).
- King, C.R. and Piatigorsky, J. Alternative RNA splicing of the murine  $\alpha$ A-crystallin gene: protein-coding information within an intron. *Cell* 32: 707-712 (1983).
- Klessing, D.F. Two Adenovirus mRNAs have a common 5' terminal leader sequence encoded at least 10kb upstream from their main coding regions. *Cell* 12: 9-21 (1977).
- Kohorn, B.D. and Rae, P.M.M. Localization of DNA sequences promoting RNA polymerase I activity in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 80: 3265-3268 (1983).
- Kornblihtt, A.R., Vibe-Pedersen, K. and Baralle, F.E. Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats. *Nucleic Acids Res.* 12: 5853-5868 (1984).



- Kramers, P.G.N., Schalet, A.P., Paradi, E. and Huiser-Hoogteyling, L. High proportion of multi-locus deletions among hycanthone-induced X-linked recessive lethals in *Drosophila melanogaster*. *Mutation Res.* 107: 187-201 (1983).
- Kress, M., Glaros, D., Khoury, G. and Jay, G. Alternative RNA splicing in expression of the H-2K gene. *Nature* 306: 602-604 (1983).
- Krieg, P.A. and Melton, D.A. Formation of the 3' end of histone mRNA by post-transcriptional processing. *Nature* 308: 203-206 (1984(a)).
- Krieg, P.A. and Melton, D.A. Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs. *Nucleic Acids Res.* 12: 7057-7070 (1984(b)).
- Kulkarni, S.J. and Padhye, A. Temperature-sensitive paralytic mutations on the second and third chromosomes of *Drosophila melanogaster*. *Genet. Res.* 40: 191-199 (1982).
- Kuroiwa, A., Hafen, E. and Gehring, W.J. Cloning and transcriptional analysis of the segmentation gene *fushi tarazu* of *Drosophila*. *Cell* 37: 825-831 (1984).
- Langridge, P. and Feix, G. A zein gene of maize is transcribed from two widely separated promotor regions. *Cell* 34: 1015-1022 (1983).
- Lassar, A.B., Martin, P.L. and Roeder, R.G. Transcription of class III genes: formation of pre-initiation complexes. *Science* 222: 740-748 (1983).
- Laughton, A. and Scott, M.P. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. *Nature* 310: 25-31 (1984).
- Lee, M.G.S., Loomis, C. and Cowan, N.J. Sequence of an expressed human  $\beta$ -tubulin gene containing ten *Alu* family members. *Nucleic Acids Res.* 12: 5823-5836 (1984).

- Lefevre, G. A photographic representation and interpretation of the polytene chromosomes of *Drosophila melanogaster* salivary glands. In *The Genetics and Biology of Drosophila*. Eds Ashburner, M. and Novitski, E. Academic Press. Vol.1a : 31-65 (1976).
- Lefevre, G. The distribution of randomly recovered X-ray induced sex-linked genetic effects in *Drosophila melanogaster*. *Genetics* 99: 461-480 (1981).
- Lehmann, R., Jimenez, F., Dietrich, U. and Campos-Ortega, J.A. On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. *Roux's Arch. Develop. Biol.* 192: 62-74 (1983).
- Levinger, L. and Varshavsky, A. Selective arrangement of ubiquitinated and D1 protein-containing nucleosomes within the *Drosophila* genome. *Cell* 28: 375-385 (1982(a)).
- Levinger, L. and Varshavsky, A. Protein D1 preferentially binds A+T-rich DNA *in vitro* and is a component of *Drosophila melanogaster* nucleosomes containing A+T-rich satellite DNA. *Proc. Natl. Acad. Sci. U.S.A.* 79: 7152-7156 (1982(b)).
- Levis, R., O'Hare, K. and Rubin, G.M. Effects of transposable element insertions on RNA encoded by the *white* gene of *Drosophila*. *Cell* 38: 471-481 (1984).
- Levy, L.S. and Manning, J.E. Messenger RNA sequence complexity and homology in developmental stages of *Drosophila*. *Dev. Biol.* 85: 1418-1419 (1981).
- Lewis, E.B. A gene complex controlling segmentation in *Drosophila*. *Nature* 276: 565-570 (1978).
- Lewis, N. and Gibson, J. Variation in amount of enzyme protein in natural populations. *Biochem. Genet.* 16: 159-170 (1978).

- Lifschytz, E. Fine-structure analysis of the chromosome recombinational patterns at the base of the X chromosome of *Drosophila melanogaster*. *Mutation Res.* 13: 35-47 (1971).
- Lifschytz, E. Fine-structure analysis and genetic organisation of the base of the X chromosome in *Drosophila melanogaster*. *Genetics* 88: 457-467 (1978).
- Lifschytz, E. Sequence replication and banding organization in the polytene chromosomes of *Drosophila melanogaster*. *J. Mol. Biol.* 164: 17-34 (1983).
- Lifschytz, E. and Falk, R. Fine structure analysis of a chromosome segment in *Drosophila melanogaster*. Analysis of X-ray induced lethals. *Mutation Res.* 6: 235-244 (1968).
- Lifschytz, E. and Falk, R. Fine structure analysis of a chromosome segment in *Drosophila melanogaster*. *Mutation Res.* 8: 147-155 (1969).
- Lifton, R.P., Goldberg, M.L., Karp, R.W. and Hogness, D.S. The organisation of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. *Cold Spring Harbor Symp. Quant. Biol.* 42: 1047-1051 (1977).
- Lindsley, D.L. and Grell, E.H. *Genetic variations of Drosophila melanogaster*. Carnegie Institute, Washington Publication. p627 (1968).
- Lis, J.T., Ish-Horowicz, D. and Pinchin, S.M. Genomic organization and transcription of the  $\alpha\beta$  heat shock DNA in *Drosophila melanogaster*. *Nucleic Acids Res.* 9: 5297-5310 (1981).
- Lis, J.T., Prestige, L. and Hogness, D.S. A novel arrangement of tandemly repeated genes at a major heat shock site in *D. melanogaster*. *Cell* 14: 901-919 (1978).
- Maniatis, T., Fritsch, E.F. and Sambrook, J. *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory. New York. (1982).

- Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. The isolation of structural genes from libraries of eukaryotic DNA. *Cell* 15: 687-701 (1978).
- Marcu, K.B. Immunoglobulin heavy chain constant-region genes. *Cell* 29: 719-721 (1982).
- Markow, T.A. and Merriam, J. Phototactic and geotactic behaviour of counter-current defective mutants of *Drosophila melanogaster*. *Behav. Genet.* 7: 447-455 (1977).
- McCarron, M., O'Donnell, J., Chovnick, A., Bhullar, B.S., Hewitt, J. and Candido, E.P.H. Organisation of the *rosy* locus in *Drosophila melanogaster*: further evidence in support of a *cis*-acting control element adjacent to the xanthine dehydrogenase structural element. *Genetics* 91: 275-293 (1979).
- McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A. and Gehring, W.J. A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* 308: 428-433 (1984).
- McGinnis, W., Shermoen, A.W. and Beckendorf, S.K. A transposable element inserted just 5' to a *Drosophila* glue protein gene alters gene expression and chromatin structure. *Cell* 34: 75-84 (1983(a)).
- McGinnis, W., Shermoen, A.W., Heemkerk, J. and Beckendorf, S.K. DNA sequence changes in an upstream DNA'ase I hypersensitive region are correlated with reduced gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 80: 1063-1067 (1983(b)).
- McMaster, G.K. and Carmichael, G.G. Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc. Natl. Acad. Sci. U.S.A.* 74: 4835-4838 (1977).

- Medford, R.M., Nguyen, H.T., Destree, A.T., Summers, E. and Nadal-Ginard, B.  
A novel mechanism of alternative RNA splicing for the developmentally regulated generation of troponin T isoforms from a single gene. *Cell* 38: 409-421 (1984).
- Melton, D.A. Injected anti-sense RNAs specifically block messenger RNA translation *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* 82: 144-148 (1985).
- Mercola, M., Wang, X.F., Olsen, J. and Calame, K. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. *Science* 221: 663-665 (1983).
- Merriam, J.R. FM7: first multiple seven. *Drosophila Information Service* 43: 64 (1968).
- Meyerowitz, E.M. and Hogness, D.S. Molecular organization of a *Drosophila* puff site that responds to ecdysone. *Cell* 28: 165-176 (1982).
- Miklos, G.L.G. The isolation of high molecular weight DNA from adult heads of *D. melanogaster*. *Drosophila Information Service* 60: 222-223 (1984).
- Miklos, G.L.G., Healy, M.J., Pain, P., Howells, A.J. and Russell, R.J. Molecular and genetic studies on the euchromatin-heterochromatin transition region of the X chromosome of *Drosophila melanogaster*. 1. A cloned entry point near to the *uncoordinated* (*unc*) locus. *Chromosoma* 89: 218-227 (1984).
- Miller, J.H. *Experiments in molecular genetics*. Cold Spring Harbor Laboratory. (1972).
- Milner, R.J., Bloom, F.E., Lai, C., Lerner, R.A. and Sutcliffe, G. Brain-specific genes have identifier sequences in their introns. *Proc. Natl. Acad. Sci. U.S.A.* 81: 713-717 (1984).
- Mindrinos, M.N., Scherer, L.J., Garcini, F.J., Kwan, H., Jacobs, K.A. and Petri, W.H. Isolation and chromosomal location of putative vitelline membrane genes in *Drosophila melanogaster*. *The EMBO Journal* 4: 147-153 (1985).

- Mirault, M.E., Southgate, R. and Delwart, E. Regulation of heat-shock genes: a DNA sequence upstream of *Drosophila hsp 70* genes is essential for their induction in monkey cells. *The EMBO Journal* 1: 1279-1285 (1982).
- Mizuno, T., Chou, M.Y. and Inouye, M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (mic RNA). *Proc. Natl. Acad. Sci. U.S.A.* 81: 1966-1970 (1984).
- Modolell, J., Bender, W. and Meselson, M. *Drosophila melanogaster* mutations suppressible by the *suppressor of hairy-wing* are insertions of a 7.3 kilobase mobile element. *Proc. Natl. Acad. Sci. U.S.A.* 80: 1678-1682 (1983).
- Monson, J.M., Natzle, J., Friedman, J. and McCarthy, B.J. Expression and novel structure of a collagen gene in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 79: 1761-1766 (1982).
- Montgomery, E.A. and Langley, C.H. Transposable elements in mendelian populations II. Distribution of three copia-like elements in a natural population of *Drosophila melanogaster*. *Genetics* 104: 473-483 (1983).
- Moore, D.D., Marks, A.R., Buckley, D.I., Kapler, G., Payvar, F. and Goodman, H.M. The first intron of the human growth hormone gene contains a binding site for glucocorticoid receptor. *Proc. Natl. Acad. Sci. U.S.A.* 82: 699-702 (1985).
- Moore, G.D., Sinclair, D.A. and Grigliatti, T.A. Histone gene multiplicity and position effect variegation in *Drosophila melanogaster*. *Genetics* 105: 327-344 (1983).
- Mortiz, Th., Edstrom, J.E. and Pongs, O. Cloning of a gene localized and expressed at the ecdysteroid regulated puff 74EF in salivary glands of *Drosophila* larvae. *The EMBO Journal* 3: 289-295 (1984).
- Mottus, R., Reeves, R. and Grigliatti, T.A. Butyrate suppression of position effect variegation in *Drosophila melanogaster*. *Mol. Gen. Genet.* 178: 465-469 (1980).



- Mukhina, L.I., Kulitchkov, V.A. and Zhimulev, I.F. Distribution of chromosome rearrangement breaks along the polytene chromosomes of *D. melanogaster*. *Drosophila Information Service* 56: 98-103 (1981).
- Murray, N.E., Brammer, W.J. and Murray, K. Lambdoid phages that simplify the recovery of *in vitro* recombinants. *Mol. Gen. Genet.* 150: 53-61 (1977).
- Muskavitch, M.A.T. and Hogness, D.S. Molecular analysis of a gene in a developmentally regulated puff of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 77: 7362-7366 (1980).
- Nabeshima, Y., Fuju-Kuriyama, Y., Muramatsu, H. and Ogata, K. Alternative transcription and two modes of splicing results in two myosin light chains from one gene. *Nature* 308: 333-338 (1984).
- Nasmyth, K. Molecular analysis of a cell lineage. *Nature* 302: 670-676 (1983).
- Natzle, J.E. and McCarthy, B.J. Regulation of *Drosophila*  $\alpha$ - and  $\beta$ -tubulin genes during development. *Devel. Biol.* 104: 187-198 (1984).
- Natzle, J.E., Monson, J.M. and McCarthy, B.J. Cytogenetic location and expression of collagen-like genes in *Drosophila*. *Nature* 296: 368-371 (1982).
- Nawa, H., Kotani, H. and Nakanishi, S. Tissue specific generation of two preprotachykinin mRNAs from one gene by alternative RNA splicing. *Nature* 312: 729-734 (1984).
- Nordheim, A. and Rich, A. Negatively supercoiled simian virus 40 DNA contains Z-DNA segments with transcriptional enhancer sequences. *Nature* 303: 674-679 (1983).
- Nussinov, R. and Lennon, G.G. Structural features are as important as sequence homologies in *Drosophila* heat shock gene upstream regions. *J. Mol. Evol.* 20: 106-110 (1984).
- Nusslein-Volhard, C. and Weischaus, E. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287: 795-801 (1980).

- O'Farrell, P.H., Kutter, E. and Nakanisbhi, M. A restriction map of the bacteriophage T4 genome. *Mol. Gen. Genet.* 179: 421-435 (1980).
- O'Hare, K., Murphy, C., Levis, R. and Rubin, G.M. DNA Sequence of the *white* locus of *Drosophila melanogaster*. *J. Mol. Biol.* 180: 437-455 (1984).
- Ouweneel, W.J. Developmental genetics of homoeosis. *Adv. in Genet.* 18: 179-248 (1976).
- Pardue, M.L., Kedes, L.H., Weinberg, E.S. and Birnstiel, M.L. Localization of sequences coding for histone messenger RNA in the chromosomes of *Drosophila melanogaster*. *Chromosoma* 63: 135-151 (1977).
- Parker, C.S. and Topol, J. A *Drosophila* RNA polymerase II transcription factor contains a promotor region specific DNA-binding activity. *Cell* 36: 357-369 (1984(a)).
- Parker, C.S. and Topol, J. A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an *hsp-70* gene. *Cell* 37: 273-283 (1984(b)).
- Parnes, J.R., Robinoson, R.R. and Seidman, J.G. Multiple mRNA species with distinct 3' termini are transcribed from the B<sub>2</sub>-microglobulin gene. *Nature* 302: 449-452 (1982).
- Peacock, W.J., Appels, R., Endow, S. and Glover, D. Chromosomal distribution of the major insert in *Drosophila melanogaster* 28S rDNA gene. *Genet. Res.* 37: 209-214 (1981).
- Peacock, W.J., Lohe, A.R., Gerlach, W.L., Dunsmuir, P., Dennis, E.S. and Appels, R. Fine structure and evolution of DNA in heterochromatin. *Cold Spring Harbour Symp. Quant. Biol.* 42: 1121-1135 (1977).
- Pelham, H.R.B. A regulatory upstream promotor element in the *Drosophila hsp 70* heat shock gene. *Cell* 30: 517-528 (1982).
- Persson, H. and Philipson, L. Regulation of Adenovirus gene expression. *Current Topics in Microbiology and Immunology* 97: 157-203 (1982).

- Pestka, S., Daugherty, B., Jung, V., Motto, K. and Pestka, R.K. Anti-mRNA: specific inhibition of translation of single mRNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* 81: 7525-7528 (1984).
- Pirrotta, V. and Brockl, Ch. Transcription of the *Drosophila white* locus and some of its mutants. *The EMBO Journal* 3: 563-568 (1984).
- Pirrotta, V., Hadfield, C., Pretorius, G.H.J. Microdissection and cloning of the *white* locus and the 3B1-3C2 region of the *Drosophila* X chromosome. *The EMBO Journal* 2: 927-934 (1983).
- Poole, S.J., Kauver, L.M., Drew, B. and Kornberg, T. The *engrailed* locus of *Drosophila*: structural analysis of an embryonic transcript. *Cell* 40: 37-43 (1985).
- Posakony, J., Benyajati, C., McDonald, J., Sullivan, D., Rowen, R. and Aguade, M. Adh: from molecular genetics to population biology. 25<sup>th</sup> Annual *Drosophila Research Conference* (1984).
- Preiss, A., Rosenberg, U.B., Kunlin, A., Seifert, E. and Jackle, H. Molecular genetics of *Kruppel*, a gene required for segmentation of the *Drosophila* embryo. *Nature* 313: 27-32 (1985).
- Procunier, J.D. and Dunn, R.J. Genetic and molecular organization of the 5S locus and mutants in *Drosophila melanogaster*. *Cell* 15: 1087-1093 (1978).
- Quinn, W.G., Sziber, P.P. and Booker, R. The *Drosophila* memory mutant *amnesiac*. *Nature* 277: 212-214 (1979).
- Reddy, P., Zehring, W.A., Wheeler, D.A., Pirrotta, V., Hadfield, C., Hall, J.C. and Rosbash, M. Molecular analysis of the *period* locus in *Drosophila melanogaster* and identification of a transcript involved in biological rhythms. *Cell* 38: 701-710 (1984).
- Richards, G., Cassab, A., Bourouis, M., Jarry, B. and Dessocis, C. The normal developmental regulation of a cloned *Sgs-3* 'glue' gene chromosomally integrated in *Drosophila melanogaster* by P element transformation. *The EMBO Journal* 12: 2137-2142 (1983).

- Riddell, D.C., Higgins, M.J., McMillan, B.J. and White, B.N. Structural analysis of three vitellogenin genes in *Drosophila melanogaster*. *Nucleic Acids Res.* 9: 1323-1338 (1981).
- Rieger, R., Michaelis, A. and Green, M.M. *A Glossary Of Genetics and Cytogenetics*. Springer-Verlag. (1968).
- Rosenberg, U.B., Preiss, A., Seifert, E. Jackle, H. and Knight, D.C. Production of phenocopies by *Kruppel* anti-sense RNA injection into *Drosophila* embryos. *Nature* 313: 703-706 (1985).
- Rosenfeld, M.G., Mermod, J.J., Amara, S.G., Swanson, L.W., Sawchenko, P.E., Rivier, J., Vale, W.W. and Evans, R.M. Production of a novel neuropeptide encoded by the calcitonin gene via tissue-specific RNA processing. *Nature* 304: 129-135 (1983).
- Rozek, C.E., and Davidson, N. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. *Cell* 32: 23-34 (1983).
- Rubin, G.M., Brorein, W.J., Dunsmuir, P., Flavell, A.J., Levis, R., Strobel, E., Toole, J.J. and Young, E. *copia*-like transposable elements in the *Drosophila* genome. *Cold Spring Harbor Symp. Quant. Biol.* 45: 619-628 (1980).
- Rubin, G.M., and Spradling, A.C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* 218: 348-353 (1982).
- Rudkin, G.T. Replication in polytene chromosomes. In *Results and Problems in Cell Differentiation* Eds. Beerman, W. Springer-Verlag. Vol. 4: 59-85 (1972).
- Rushlow, C.A., Bender, W. and Chovnick, A. Studies on the mechanism of heterochromatic position effect at the *rosy* locus of *Drosophila melanogaster*. *Genetics* 108: 603-615 (1984).

- Ryffel, G.U., Muellener, D.B., Wyler, T., Wahli, W. and Weber, R. Transcription of single copy vitellogenin gene of *Xenopus* involves expression of middle repetitive DNA. *Nature* 291: 429-431 (1981).
- Saigo, K., Millstein, L. and Thomas, C.A. Jr. The organization of *Drosophila melanogaster* histone genes. *Cold Spring Harbor Symp. Quant. Biol.* 45: 815-827 (1980).
- Salkoff, L. Genetic and voltage clamp analysis of a *Drosophila* potassium channel. *Cold Spring Harbor Symp. Quant. Biol.* 48: 221-231 (1983).
- Salkoff, L. and Wyman, R. Genetic modification of potassium channels in *Drosophila Shaker* mutants. *Nature* 293: 228-230 (1981).
- Sampsell, B. Isolation and genetic characterization of alcohol dehydrogenase thermostability variants occurring in natural populations of *Drosophila melanogaster*. *Biochem. Genet.* 15: 971-988 (1977).
- Savakis, C., Demetri, G. and Cherbas, P. Ecdysteroid-inducible polypeptides in a *Drosophila* cell line. *Cell* 22: 665-674 (1980).
- Savakis, C., Kaehler, M.M.D. and Cherbas, P. cDNA clones for the ecdysone-inducible polypeptide (EIP) mRNAs of *Drosophila* Kc cells. *The EMBO Journal* 3: 235-243 (1984).
- Schalet, A. Crossing over in the major heterochromatic region of the X chromosome in normal and inverted sequences. *Drosophila Information Service* 48: 111-113 (1972).
- Schalet, A. and Lefevre, G. Jr. The localization of 'ordinary' sex-linked genes in section 20 of the polytene X chromosome of *Drosophila melanogaster*. *Chromosoma* 44: 183-202 (1973).
- Schalet, A. and Lefevre, G. Jr. The proximal region of the X chromosome. In *The Genetics and Biology of Drosophila*. Eds. Ashburner, M. and Novitski, E. Academic Press. Vol. 1b: 847-902 (1976).

- Schalet, A. and Roberts, P.A. *arth*: arthritic. *Drosophila Information Service* 50: 23 (1973).
- Schimke, R.T. Gene amplification in cultured animal cells. *Cell* 37: 705-713 (1984).
- Schmidt-Nielsen, B. and Hall, L.M. An abnormal walking mutant associated with a translocation. *Drosophila Information Service* 52: 71-72 (1977).
- Scholnick, S.B., Morgan, B.A. and Hirsh, J. The cloned dopa decarboxylase gene is developmentally regulated when reintegrated into the *Drosophila* genome. *Cell* 34: 37-45 (1983).
- Schultz, J. Variegation in *Drosophila* and the inert chromosome regions. *Proc. Natl. Acad. Sci.* 22: 27-33 (1936).
- Schwartz, H.E., Lockett, T.J. and Young, M.W. Analysis of transcripts from two families of nomadic DNA. *J. Mol. Biol.* 157: 49-68 (1982).
- Schwarzbauer, J.E., Tamkun, J.W., Lemischka, I.R. and Hynes, R.O. Three different fibronectin mRNAs arise by alternative splicing within the coding region. *Cell* 35: 421-431 (1983).
- Scott, M.P. and Weiner, A.J. Structural relationships among genes that control development: sequence homology between the *Antennapedia*, *Ultrabithorax* and *fushi tarazu* loci of *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 81: 4115-4119 (1984).
- Scott, M.P., Weiner, A.J., Hazelrigg, T.I., Polisky, B.A., Pirrotta, V. Scalenghe, F. and Kaufman, T.C. The molecular organization of the *Antennapedia* locus of *Drosophila*. *Cell* 35: 763-776 (1983).
- Segraves, W.A., Louis, C., Schedl, P. and Jarry, B.P. Isolation of the rudimentary locus of *Drosophila melanogaster*. *Mol. Gen. Genet.* 189: 34-40 (1983).
- Segraves, W.A., Louis, C., Tsubota, S., Schedl, P., Rawls, J.M. and Jarry, B.P. The rudimentary locus of *Drosophila melanogaster*. *J. Mol. Biol.* 175: 1-17 (1984).



- Shermoen, A.W. and Beckendorf, S.K. A complex of interacting DNA'ase 1 hypersensitive sites near the *Drosophila* glue protein gene, *Sgs-4*. *Cell* 29: 601-607 (1982).
- Simpson, P.R. Molecular studies of a dispersed simple DNA sequence in *Drosophila melanogaster*. Ph.D. Thesis. A.N.U. (1984).
- Smith, D.F., McClelland, A., White, B.N., Addison, C.F. and Glover D.M. The molecular cloning of a dispersed set of developmentally regulated genes which encode the major larval serum protein of *Drosophila*. *Cell* 23: 441-449 (1981).
- Smith, G.P. Evolution of repeated DNA sequences by unequal crossover. *Science* 191: 528-535 (1976).
- Snyder, M., Hirsh, J. and Davidson, N. The cuticle genes of *Drosophila* : a developmentally regulated gene cluster. *Cell* 25: 167-177 (1981).
- Snyder, M., Hunkapiller, M., Yuen, D., Silvert, D., Fristrom, J. and Davidson, N. Cuticle protein genes of *Drosophila*: structure, organization and evolution of four clustered genes. *Cell* 29: 1027-1040 (1982).
- Snyder, M., Kimbrell, D., Hunkapiller, M., Hill, R., Fristrom, J. and Davidson, N. A transposable element that splits the promotor region inactivates a *Drosophila* cuticle protein gene. *Proc. Natl. Acad. Sci. U.S.A.* 79: 7430-7434 (1982).
- Soberon, X., Covarrubias, L. and Bolivar, F. Construction and characterization of new cloning vehicles IV. Deletion derivatives of pBR322 and pBR325. *Gene* 9: 287-305 (1980).
- Sodroski, J.G., Rosen, C.A. and Haseltine, W.A. Trans-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. *Science* 225: 381-385 (1984).
- Sondergaard, L. Dominant cold paralytic mutations on the X-chromosome of *Drosophila melanogaster*. *Hereditas* 90: 93-101 (1979).

- Southern, E.M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517 (1975).
- Spierer, P., Spierer, A., Bender, W. and Hogness, D.S. Molecular mapping of genetic and chromomeric units in *Drosophila melanogaster*. *J. Mol. Biol.* 168: 35-50 (1983).
- Spofford, J.B. Position effect variegation in *Drosophila*. In *The Genetics and Biology of Drosophila*. Eds. Ashburner, M. and Novitski, E. Academic Press. Vol. 1c: 955-1018 (1976).
- Spradling, A.C. The organization and amplification of two chromosomal domains containing *Drosophila* chorion genes. *Cell* 27: 193-201 (1981).
- Spradling, A.C. and Mahowald, A.P. Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 77: 1096-1100 (1980).
- Spradling, A.C., Rigan, M.E., Mahowald, A.P., Scott, M. and Craig, E.A. Two clusters of genes for major chorion proteins of *Drosophila melanogaster*. *Cell* 19: 905-914 (1980).
- Spradling, A.C. and Rubin, G.M. *Drosophila* genome organisation: conserved and dynamic aspects. *Ann. Rev. Genetics* 15: 219-264 (1981).
- Spradling, A.C. and Rubin, G.M. Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* 218: 341-347 (1982).
- Spradling, A.C. and Rubin, G.M. The effect of chromosomal position on the expression of the *Drosophila* xanthine dehydrogenase gene. *Cell* 34: 47-57 (1983).
- Steffensen, D.M., Appels, R. and Peacock, W.J. The distribution of two highly repeated DNA sequences within *Drosophila melanogaster* chromosomes. *Chromosoma* 82: 525-541 (1981).
- Stellar, H. and Pirrotta, V. A transposable P vector that confers selectable G418 resistance to *Drosophila* larvae. *The EMBO Journal* 4: 167-171 (1985).

- Steward, R., McNally, F.J. and Schedl, P. Isolation of the *dorsal* locus of *Drosophila*. *Nature* 311: 262-265 (1984).
- Strauss, F. and Varshavsky, A. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* 37: 889-901 (1984).
- Strobel, E. Mobile dispersed repeated DNA elements in the *Drosophila* genome. *Fed. Proc.* 41: 2656-2658 (1982).
- Strobel, E., Dunsmuir, P. and Rubin, G.M. Polymorphisms in the chromosomal locations of elements of the 412, *copia* and 297 dispersed repeated gene families in *Drosophila*. *Cell* 17: 429-439 (1979).
- Sutcliffe, J.G., Milner, R.J. and Bloom, F.E. Cellular localization and function of the proteins encoded by brain-specific mRNAs. *Cold Spring Harbor Symp. Quant. Biol.* 48: 477-484 (1983).
- Swaroop, A., Paco-Larson, M.L. and Garen, A. Molecular genetics of a transposon-induced dominant mutation in the *Drosophila* locus *Glued*. *Proc. Natl. Acad. Sci. U.S.A.* 82: 1751-1755 (1985).
- Takeda, Y., Ohlendorf, D.H., Anderson, W.F. and Matthews, B.W. DNA binding proteins. *Science* 221: 1020-1026 (1983).
- Tartof, K.D., Hobbs, C. and Jones, M. A structural basis for variegating position effects. *Cell* 37: 869-878 (1984).
- Tchurikov, N.A., Naumova, A.K., Zelentsova, E.S. and Georgiev, G.P. A cloned unique gene of *Drosophila melanogaster* contains a repetitive 3' exon whose sequence is present at the 3' ends of many different mRNAs. *Cell* 28: 365-373 (1982).
- Tempel, B.L., Livingstone, M.S. and Quinn, W.G. Mutations in the dopa decarboxylase gene affect learning in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 81: 3577-3581 (1984).

- Thomas, J.B. and Wyman, R.J. A mutation in *Drosophila* alters normal connectivity between two identified neurones. *Nature* 298: 650-651 (1982).
- Thomas, P.S. Hybridisation of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. U.S.A.* 77: 5201-5205 (1980).
- Tobin, S.L., Zulauf, E., Sanchez, F., Craig, E.A. and McCarthy, B.J. Multiple actin-related sequences in the *Drosophila melanogaster* genome. *Cell* 19: 121-131 (1980).
- Truett, M.A., Jones, R.S. and Potter, S.S. Unusual structure of the FB family of transposable elements in *Drosophila*. *Cell*: 24: 753-763 (1981).
- Ullu, E. and Tschudi, C. *Alu* sequences are processed 7S RNA genes. *Nature* 312: 171-172 (1984).
- Van Ness, J., Maxwell, I.A. and Hahn, W.E. Complex population of nonpolyadenylated messenger RNA in mouse brain. *Cell* 18: 1341-1349 (1979).
- Vaslet, C.A., O'Connell, P., Izquierdo, M. and Rosbash, M. Isolation and mapping of a cloned ribosomal protein gene of *Drosophila melanogaster*. *Nature* 285: 674-676 (1980).
- Velissariou, V. and Ashburner, M. Cytogenetic and genetic mapping of a salivary gland secretion protein in *Drosophila melanogaster*. *Chromosoma* 84: 173-185 (1981).
- Vincent, A., O'Connell, P., Gray, M.R. and Rosbash, M. *Drosophila* maternal and embryo mRNAs transcribed from a single transcription unit use alternate combination of exons. *The EMBO Journal* 3: 1003-1013 (1984).
- Walldorf, U., Richter, S., Ryseck, R.P., Steller, H., Edstrom, J.E., Bautz, E.K.F. and Havemann, B. Cloning of heat shock locus 93D from *Drosophila melanogaster*. *The EMBO Journal* 3: 2499-2504 (1984).

- Weideli, H., Schedl, P., Artavanis-Tsakonas, S., Steward, R., Yuan, R. and Gehring, W.J. Purification of a protein from unfertilized eggs of *Drosophila* with specific affinity for a defined DNA sequence and the cloning of this DNA sequence in bacterial plasmids. *Cold Spring Harbor Symp. Quant. Biol.* 42: 693-700 (1977).
- Weiner, A.J., Scott, M.P. and Kaufman, T.C. A molecular analysis of *fushi tarazu*, a gene in *Drosophila melanogaster* that encodes a product affecting embryonic segment number and cell fate. *Cell* 37: 843-851 (1984).
- Weisbrod, S. Active chromatin. *Nature* 297: 289-295 (1982).
- Wensink, P.C., Finnegan, D.J., Donelson, J. and Hogness, D.S. A system for mapping DNA sequences in the chromosomes of *Drosophila melanogaster*. *Cell* 3: 315-325 (1974).
- Wensink, P.C., Tabata, S. and Pachl, C. The clustered and scrambled arrangement of moderately repetitive elements in *Drosophila* DNA. *Cell* 18: 1231-1246 (1979).
- Wharton, K.A., Yedvobnick, B., Finnerty, V.G. and Artavanis-Tsakonas, S. *opa*, a novel family of transcribed repeats shared by the *Notch* locus and other developmentally regulated loci in *Drosophila melanogaster*. *Cell* 40: 55-62 (1985).
- White, R.A.H. and Wilcox, M. Protein products of the bithorax complex in *Drosophila*. *Cell* 39: 163-171 (1984).
- Whitfield, P.L., Seeburg, P.H. and Shine, J. The human pro-opiomelanocortin gene: organization, sequence and interspersions with repetitive DNA. *DNA* 1: 133-143 (1982).
- Wieschaus, E., Nusslein-Volhard, C. and Jurgens, G. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. *Roux's Arch. Devel. Biol.* 193: 296-307 (1984).



- Will, H. and Bautz, E.K.F. Immunological identification of a chromocenter-associated protein in polytene chromosomes of *Drosophila*. *Exp. Cell. Res.* 125: 401-410 (1980).
- Williamson, J.H. The genetics of the Y chromosome. In *The Genetics and Biology of Drosophila*. Eds. Ashburner, M. and Novitski, E. Academic Press. Vol. 1b: 667-699 (1976).
- Woolf, C.M. and Passage, M.B. Genetic variability of the tumorous-head maternal effect in *Drosophila melanogaster*. *Mol. Gen. Genet.* 178: 423-427 (1980).
- Wright, S., Rosenthal, A., Flavell, R. and Grosveld, F. DNA sequences required for regulated expression of  $\beta$ -globin genes in murine erythroleukemia cells. *Cell* 38:265-273 (1984).
- Wu, C. The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNA'ase I. *Nature* 286: 854-860 (1980).
- Wu, C. Two protein-binding sites in chromatin implicated in the activation of heat-shock genes. *Nature* 309: 229-234 (1984(a)).
- Wu, C. Activating protein factor binds *in vitro* to upstream control sequences in heat shock chromatin. *Nature* 311: 81-84 (1984).
- Wu, C.F. and Ganetzky, B. Genetic alteration of nerve membrane excitability in temperature-sensitive paralytic mutants of *Drosophila melanogaster*. *Nature* 286: 814-816 (1980).
- Wu, C.F., Ganetzky, B., Jan, L.Y., Jan, Y.N. and Benzer, S. A *Drosophila* mutant with a temperature-sensitive block in nerve conduction. *Proc. Natl. Acad. Sci. U.S.A.* 75: 4047-4051 (1978).
- Wyman, R.J. and Thomas, J.B. What genes are necessary to make an identified synapse? *Cold Spring Harbor Symp. Quant. Biol.* 48: 641-652 (1983).
- Young, B.S., Pession, A., Traverse, K.L., French, C. and Pardue, M.L. Telomere regions in *Drosophila* share complex DNA sequences with pericentric heterochromatin. *Cell* 34: 85-94 (1983).



- Young, M.W., and Schwartz, H.E. Nomadic gene families in *Drosophila*. *Cold Spring Harbor Symp. Quant. Biol.* 45: 629-640 (1980).
- Young, R.A. and Davis, R.W. Efficient isolation of genes using antibody probes. *Proc. Natl. Acad. Sci. U.S.A.* 80: 1194-1198 (1983).
- Young, R.A., Hagenbuchle, O. and Schibler, U. A single mouse  $\alpha$ -amylase gene specifies two different tissue-specific mRNAs. *Cell* 23: 451-458 (1981).
- Yund, M.A. and Germeraad, S. *Drosophila* Development. In *Molecular Biology* Eds. Leighton, T. and Loomis, W.F. Academic Press. pp237-360 (1980).
- Zachar, Z. and Bingham, P.M. Regulation of *white* locus expression: the structure of mutant alleles at the *white* locus of *Drosophila melanogaster*. *Cell* 30: 529-541 (1982).
- Zachau, H.G. and Igo-Kemenes, T. Face to phase with nucleosomes. *Cell* 24: 597-598 (1981).
- Zaret, K.S. and Yamamoto, K.R. Reversible and persistent changes in chromatin structure accompany activation of a glucocorticoid-dependent enhancer element. *Cell* 38: 29-38 (1984).
- Zehner, Z.E. and Paterson, B.M. Characterization of the chicken vimentin gene: single copy gene producing multiple mRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 80: 911-915 (1983).
- Zhimulev, I.F., Semeshin, V.F., Kulichkov, V.A. and Belyaeva, E.S. Intercalary heterochromatin in *Drosophila*. 1. Location and general characteristics. *Chromosoma* 87: 197-228 (1982).
- Ziff, E.B. Transcription and RNA processing by the DNA tumour viruses. *Nature* 287: 491-499 (1980).
- Zimmerman, J.L., Fouts, D.L., Levy, L.S. and Manning, J.E. Nonadenylated mRNA is present as polyadenylated RNA in nuclei of *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 79: 3148-3152 (1982).
- Zimmerman, J.L., Fouts, D.L. and Manning, J.E. Evidence for a complex class of nonadenylated mRNA in *Drosophila*. *Genetics* 95: 673-691 (1980).